Lower bounds in the convolution structure density model

Oleg Lepski, Thomas Willer

Lower bounds in the convolution structure density model

O.V. Lepski * **T. Willer** *

Institut de Mathématique de Marseille
Aix-Marseille Université
39, rue F. Joliot-Curie
13453 Marseille, France
e-mail: oleg.lepski@univ-amu.fr
e-mail: thomas.willer@univ-amu.fr

Abstract:
The aim of the paper is to establish asymptotic lower bounds for the minimax risk in two generalized forms of the density deconvolution problem. The observation consists of an independent and identically distributed (*i.i.d.*) sample of n random vectors in $\mathbb{R}^d$. Their common probability distribution function $\mathfrak{p}$ can be written as $\mathfrak{p} = (1-\alpha)f + \alpha[f \star g]$, where f is the unknown function to be estimated, g is a known function, α is a known proportion, and $\star$ denotes the convolution product. The bounds on the risk are established in a very general minimax setting and for moderately ill posed convolutions. Our results show notably that neither the ill-posedness nor the proportion α play any role in the bounds whenever $\alpha \in [0,1)$, and that a particular inconsistency zone appears for some values of the parameters. Moreover, we introduce an additional boundedness condition on f and we show that the inconsistency zone then disappears.

AMS 2000 subject classifications: 62G05, 62G20.
Keywords and phrases: generalized deconvolution model, density estimation, adaptive estimation, Nikol'skii spaces, $\mathbb{L}_p$–risk, minimax rates.

1. Introduction.

In statistics, deconvolution is one of the most well known model among non parametric ill-posed inverse problems, where basically one wishes to recover an unknown real-valued function f after it has been smoothed by a known operator. We will consider a multidimensional setting where this operator is the convolution one, defined via some integrable real-valued "noise" function g.

Generalized deconvolution model. Our first model consists in a mix between the direct observation problem and the pure deconvolution problem. Consider the following observation scheme:

$$Z_i = X_i + \epsilon_i Y_i, \quad i = 1, \dots, n, \tag{1.1}$$

where $X_i, i = 1, \dots, n$, are *i.i.d.* d-variate random vectors with common density f to be estimated. The noise variables $Y_i, i = 1, \dots, n$, are *i.i.d.* d-variate random vectors with known common density g. At last $\varepsilon_i \in \{0,1\}, i = 1, \dots, n$, are *i.i.d.* Bernoulli random variables with $\mathbb{P}(\varepsilon_1 = 1) = \alpha$, where $\alpha \in [0,1]$ is supposed to be known.

Let us note that the case $\alpha = 1$ corresponds to the pure deconvolution model $Z_i = X_i + Y_i,\ i = 1, \dots, n$, whereas the case $\alpha = 0$ corresponds to the direct observation scheme $Z_i = X_i,\ i =$

*This work has been carried out in the framework of the Labex Archimède (ANR-11-LABX-0033) and of the A*MIDEX project (ANR-11-IDEX-0001-02), funded by the "Investissements d'Avenir" French Government program managed by the French National Research Agency (ANR).

$1, \ldots, n$. The intermediate case $\alpha \in (0,1)$ can be treated as the mathematical modeling of the following situation. One part of the data, namely $(1-\alpha)n$, is observed without noise. If the indexes corresponding to these observations were known, the density f could be estimated using only this part of the data, with the accuracy corresponding to the direct case. The main question we will address in this intermediate case is whether the same accuracy would be achievable if the latter information is not available? We will see that the answer is positive, but the construction of optimal estimation procedures is based upon ideas corresponding to the pure deconvolution model.

Convolution structure density model. Our second model is a generalization of the generalized deconvolution model itself. For any fixed $\alpha \in [0,1]$ and $R > 1$, and for any $g \in \mathbb{L}_1\big(\mathbb{R}^d\big)$ introduce

$$\mathbb{F}_g(R) = \Big\{ f \in \mathbb{B}_{1,d}(R) : (1-\alpha)f + \alpha[f \star g] \in \mathfrak{P}\big(\mathbb{R}^d\big) \Big\},$$

where $\mathbb{B}_{1,d}(R)$ denotes the ball of radius R in $\mathbb{L}_1\big(\mathbb{R}^d\big)$ and $\mathfrak{P}\big(\mathbb{R}^d\big)$ denotes the set of probability densities on $\mathbb{R}^d$. Suppose that we observe *i.i.d.* vectors $Z_i, i = 1, \ldots, n$, with common unknown density $\mathfrak{p}$ satisfying the following structural assumption:

$$\mathfrak{p} = (1-\alpha)f + \alpha[f \star g], \quad f \in \mathbb{F}_g(R), \;\; \alpha \in [0,1]. \tag{1.2}$$

Here $\alpha \in [0,1]$ and $g : \mathbb{R}^d \to \mathbb{R}$ are supposed to be known and $f : \mathbb{R}^d \to \mathbb{R}$ is the function to be estimated. We remark that the observations scheme (1.2) coincides with the observations scheme (1.1) if additionally $f, g \in \mathfrak{P}\big(\mathbb{R}^d\big)$.

What is more, the convolution structure density model (1.2) can be studied for an arbitrary $g \in \mathbb{L}_1\big(\mathbb{R}^d\big)$ and $f \in \mathbb{F}_g(R)$, but later on we will restrict ourselves to the case $g \in \mathfrak{P}\big(\mathbb{R}^d\big)$. Then on the one hand we have $\mathfrak{P}\big(\mathbb{R}^d\big) \subset \mathbb{F}_g(R)$ for any $R > 1$, but reciprocally some function $f \in \mathbb{F}_g(R)$ is not necessarily a density, except in the particular case $\alpha = 0$.

As it stands, one may question the reason for the introduction of the *convolution structure density model.* In fact, the main motivation lies in the calculation of the upper bounds of the minimax risk defined further. Indeed, in a separate work, the authors have constructed a performant estimation technique, along the lines of what Goldenshluger and Lepski (2014) developed in the direct case. However this estimator, like (to the authors knowledge) most of those available in the literature, exploit neither $f \geq 0$ nor $\int f = 1$. Thus it is interesting to consider a broader setting than the generalized deconvolution model, by removing the assumption that f is a probability density, and to check if some estimators of interest are optimal in this new setting.

The minimax framework. We will study the deconvolution problem in the minimax perspective, which is very widespread in the statistical literature (see Tsybakov (2009) for examples of problems and methods in this setting). We want to estimate the target function f using the observations $Z^{(n)} = (Z_1, \ldots, Z_n)$ given in each one of the two models (1.1) and (1.2). By estimator we mean any $Z^{(n)}$-measurable map $\hat{f} : \mathbb{R}^n \to \mathbb{L}_p\big(\mathbb{R}^d\big)$.

The accuracy of an estimator $\hat{f}$ is measured by the $\mathbb{L}_p$–risk

$$\mathcal{R}_p^{(n)}[\hat{f}, f] := \Big(\mathbb{E}_f \|\hat{f} - f\|_p^p\Big)^{1/p}, \; p \in [1, \infty), \quad \mathcal{R}_\infty^{(n)}[\hat{f}, f] := \mathbb{E}_f\big(\|\hat{f} - f\|_\infty\big).$$

Here $\mathbb{E}_f$ denotes the expectation with respect to the probability measure $\mathbb{P}_f$ of the observations $Z^{(n)} = (Z_1, \ldots, Z_n)$, $\|\cdot\|_p$, $p \in [1, \infty)$ denotes the $\mathbb{L}_p$-norm on $\mathbb{R}^d$, and $\|\cdot\|_\infty$ denotes the supremum norm on $\mathbb{R}^d$. The objective is to construct an estimator of f possessing a small $\mathbb{L}_p$–risk.

In the framework of the minimax approach, the density f is assumed to belong to a functional class Σ, which is specified on the basis of some prior information on f. Then a natural accuracy measure of an estimator $\hat{f}$ is its maximal $\mathbb{L}_p$–risk over Σ,

$$\mathcal{R}_p^{(n)}[\hat{f};\Sigma] = \sup_{f\in\Sigma} \mathcal{R}_p^{(n)}[\hat{f}, f].$$

The main question is: how to construct a *rate–optimal*, or *optimal in order*, estimator $\hat{f}_*$ such that

$$\mathcal{R}_p^{(n)}[\hat{f}_*;\Sigma] \asymp \phi_n(\Sigma) := \inf_{\hat{f}} \mathcal{R}_p^{(n)}[\hat{f};\Sigma], \quad n\to\infty?$$

Here the infimum is taken over all possible estimators, and $\phi_n(\Sigma)$ is called the minimax risk. We refer to the outlined problem as the *problem of minimax density deconvolution/estimation with $\mathbb{L}_p$–loss on the class* Σ. Then the first problem arising here consists in bounding from below the minimax risk $\phi_n(\Sigma)$ in an optimal way, i.e. with bounds as large as possible as for the dependency on n (and on other parameters appearing in Σ if possible). The aim of the paper is to address this problem, whereas the determination of the upper bounds and the development of rate-optimal estimators will be addressed in a separate paper.

In the framework of the *convolution structure density model* we will consider $\Sigma = \mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L}\big)\cap\mathbb{F}_g(R)$, where $\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L}\big)$ is the anisotropic Nikol'skii class defined further. Here we only note that a Nikol'skii class belongs to the family of Besov functional classes and it is, in fact, the largest Besov class, see Nikol'skii (1977), chapter 6.2.

Moreover we will consider $\Sigma = \mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L}\big)\cap\mathfrak{P}\big(\mathbb{R}^d\big)$, which allows to get the results in the framework of the *generalized deconvolution model.*

At last we study $\Sigma = \mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L},M\big)\cap\mathfrak{P}\big(\mathbb{R}^d\big)$, where

$$\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L},M\big) = \mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L}\big)\cap\big\{G:\mathbb{R}^d\to\mathbb{R}:\ \ \|G\|_\infty\le M\big\},\ \ M>0.$$

It will allow us to understand how the boundedness assumption on the underlying function may have effects on the minimax rate of convergence.

Moreover let us note that our main purpose is to precise how the lower bounds depend on the size n of the sample, when $n\to+\infty$. Furthermore we also study the dependence of the remaining constants on $\vec{L}$, which, remind, is the radium vector in the Nikol'skii space.

Overview of the existing results. As already mentioned, our models include as particular cases density estimation in the direct observation and in the pure deconvolution settings. As far as we know, the intermediate case mixing both situations has not been studied yet from a minimax perspective. However, of course there is a vast literature dealing with each one of the two particular cases.

First let us consider the direct observation setting. The latest results in the density estimation over $\Sigma = \mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L},M\big)$ under $\mathbb{L}_p$-loss, $1\le p<+\infty$, were obtained in Goldenshluger and Lepski (2014). The authors obtain the following lower and upper bounds for the minimax risk:

$$c(\ln n)^a n^{-\rho}\le\phi_n(\Sigma)\le C(\ln n)^b n^{-\rho},$$

where $\rho>0$, $a\ge0$, $b\ge0$ are constants completely determined by the class parameters $\vec{\beta}$, $\vec{r}$ and the norm index p. Different relation between these parameters leads to four regimes of the decay of minimax risk (cf. Theorem 2)and it is worth noting that $a=b$ in two zones among four. Moreover in the same paper, an inconsistency result is also obtained for the sup-norm loss. Along with the results of Lepski (2013), this gives a complete picture on density estimation is sup-norm loss.

Now let us focus on the results in the pure deconvolution setting. We only mention here papers dealing partly or completely with our framework, i.e. deconvolution in the density setting, in $\mathbb{L}_p$-loss, from a minimax perspective. Moreover, the behavior of the Fourier transform of the noise function g plays an important role in our results, as actually in all the works on deconvolution. Indeed ill-posed problems correspond to Fourier transforms decaying towards zero. Our results will be established for "moderately" ill posed problem, so we detail only results in papers studying that type of operators. This assumption means that there exist $\vec{\mu} = (\mu_1, \dots, \mu_d) \in (0,\infty)^d$ and $\Upsilon > 0$ such that the Fourier transform $\check{g}$ of g satisfies:

$$\Upsilon_1 \prod_{j=1}^{d} (1+t_j^2)^{-\frac{\mu_j}{2}} \leq \big|\check{g}(t)\big| \leq \Upsilon_2 \prod_{j=1}^{d} (1+t_j^2)^{-\frac{\mu_j}{2}}, \quad \forall t = (t_1, \dots, t_d) \in \mathbb{R}^d.$$

Let us first present some results in dimension one. Later on $\mathbb{B}\big(\beta, r, q, L\big)$ denotes the Besov functional classes. In particular Nikol'skii class correspond to $q = +\infty$, including Holder case $r = q = +\infty$, and Sobolev class corresponds to $q = r$. Early works include in particular Devroye (1989), who developed consistent estimators under $\mathbb{L}_1$-loss (but not yet in a minimax sense) whenever the set of points where $\big|\check{g}(t)\big| \neq 0$ is of full Lebesgue measure. Later on Stefanski and Carroll (1990) establish upper and lower bounds for the $\mathbb{L}_2$ loss of a family of estimators in three cases of noise functions: normal, Cauchy and double-exponential. However here again they did not use the minimax framework yet in the choice of the regularity space for the target.

Fan (1993) uses the $\mathbb{L}_p$-loss (with $1 \leq p < +\infty$) and the Holder class $\Sigma = \mathbb{B}\big(\beta, \infty, \infty, L\big)$, $\beta > 0, L > 0$. The author also assumes that the target functions are compactly supported in some interval $[a, b]$ and obtains the following lower bound:

$$\phi_n(\Sigma) \geq C n^{-\frac{1}{2+\frac{2\mu+1}{\beta}}}. \tag{1.3}$$

This rate is obtained in several papers as a convergence rate of an estimator, and hence upper bounds for the minimax risk. For example, the $\mathbb{L}_2$ loss and the Sobolev class $\Sigma = \mathbb{B}\big(\beta, 2, 2, L\big)$ was considered by Pensky and Vidakovic (1999), Comte and al. (2006) and Hall and Meister (2007) (who made the additional condition that $\beta > \frac{1}{2}$). They all developed estimators, based on wavelets for the first paper, on penalization and model selection in the second paper, and on a ridge method in the third paper. They all establish the convergence rate of the estimators, which turns out to be (1.3). Note also that one chapter of the book by Meister (2009) gathers results on one-dimensional density estimation in $\mathbb{L}_2$ loss over Sobolev or Holder spaces. These results are similar to the aforementioned ones: lower and upper bounds are established for the minimax risk, estimators based on ridge methods, wavelets, or kernels are developed. Fan and Koo (2002) use the $\mathbb{L}_2$-loss and the Besov class $\Sigma = \mathbb{B}\big(\beta, r, q, L\big)$. They establish a lower bound for the minimax risk for any $1 \leq q \leq \infty$ under the condition $\beta > \frac{1}{r}$. Once again the rate is the same as in (1.3). They also establish an upper bound under the additional assumptions that $1 \leq r \leq 2$, $\mu < \frac{r}{2-r}(\beta + \frac{1}{2} - \frac{1}{r})$ and f has a fixed compact support.

Lounici and Nickl (2011) examined the case of the $\mathbb{L}_\infty$-loss. They obtain a lower and an upper bound for the minimax risk over the class of Holder spaces $\Sigma = \mathbb{B}\big(\beta, \infty, \infty, L\big)$ with $\beta > 0$, but they remark that the results can be generalized to the class of Besov spaces $\Sigma = \mathbb{B}\big(\beta, r, q, L\big)$. The found rate coincides with whose given in (1.3).

There are very few results in the multidimensional setting. It seems that Masry (1993) was the first paper where the deconvolution problem was studied for multivariate densities. It is worth noting that Masry (1993) considered more general weakly dependent observations and this paper formally does not deal with minimax setting. However the results obtained in this paper could be formally compared with the estimation under $\mathbb{L}_\infty$-loss ($p = \infty$) over the isotropic Hölder class of regularity 2, i.e. the Nikol'skii class with $\vec{\beta} = (2, \dots, 2)$ and $\vec{r} = (\infty, \dots, \infty)$. Let us also remark that the paper does not contain any lower bound result.

In Comte and Lacour (2013) the authors consider the $\mathbb{L}_2$-loss and the Sobolev class $\Sigma = \mathbb{B}(\vec{\beta}, \vec{2}, \vec{2}, \vec{L})$ with the restriction $\forall j \in \{1, \dots, d\},\ \beta_j > \frac{1}{2}$. They also assume that all the components of the noise vectors are independent. They then obtain the following lower bound:

$$\phi_n(\Sigma) \geq Cn^{-\frac{1}{2+\sum_{j=1}^{d} \frac{2\mu_j+1}{\beta_j}}}.$$

They also obtain an upper bound containing the same rate in n, either on the Sobolev class as above intersected with $\mathbb{L}_1$, or on the Nikol'skii class $\Sigma = \mathbb{N}_{\vec{2},d}(\vec{\beta}, \vec{L}, M)$ with the restriction $\forall j \in \{1, \dots, d\},\ \beta_j > \frac{1}{2}$.

Let us briefly mention without details other types of deconvolution problems which are beyond the scope of this paper. First there are several results on pointwise estimation, see for example Carroll and Hall (1988), Fan (1991), Goldenshluger (1999), Butucea and Tsybakov (2008a), or the already cited paper Comte and Lacour (2013). Secondly, note that other types of ill-posedness can appear. In particular, an exponential type of decay for $|\check{g}|$ yields a "severely" ill-posed problem. All of the aforementioned papers from 1993 onward studied these kinds of problems. Moreover, papers such as Stefanski (1990), Butucea and Tsybakov (2008a) and Butucea and Tsybakov (2008b) focus exclusively on severely ill-posed problems. Still another perspective is used in Butucea (2004), where the operator is moderately ill-posed but the target function is "exponentially" smooth. Note that some convolution operators are neither moderately nor severely ill-posed. For example Hall and Meister (2007) investigate the case of oscillating behavior of $|\check{g}|$, and Johnstone and Raimondo (2004) investigate the case of the boxcar deconvolution with badly approximate width (but in the white noise setting), where $|\check{g}|$ behaves in an unstable way.

Note that there is also a vast literature on deconvolution in the white noise instead of the density setting. Here one assumes that the function $f \star g$ is observed after being corrupted by an additive gaussian white noise. Then let us mention in particular the work by Johnstone et al. (2004) who assume that f and g are one dimensional and periodic and develop an adaptive estimator, which turns out to be rate-optimal as shown by the minimax bounds in Willer (2006).

The remainder of the paper is organized as follows. Section 2 describes the assumptions on the functions f and g appearing in our models. Section 3 contains the main results, which comprehend lower bounds for the minimax risk in three situations: the generalized deconvolution problem, the convolution structure model, and the generalized deconvolution problem with a uniformly bounded target density. Moreover the results in the case of the $\mathbb{L}_p$-loss and of the uniform loss are presented separately. In Section 4, we discuss the optimality of the results presented in Section 3. We present a general conjecture related to the construction of an estimation procedure which would attain the asymptotics found in Theorem 1– 3. Moreover we consider some particular cases where the construction of an estimator attaining our lower bounds is straightforward. Sections 5 and 6 contain the proofs of the main results. The proofs of auxiliary lemmas are postponed to Appendix.

2. Definitions and assumptions.

In this section we give the definition of functional classes used in the description of the minimax risk and present the assumptions imposed on the function g used in the definition of the convolution operator.

Anisotropic Nikol'skii classes. Let $(\mathbf{e}_1,\ldots,\mathbf{e}_d)$ denote the canonical basis of $\mathbb{R}^d$. For any function $f:\mathbb{R}^d\to\mathbb{R}^1$ and any real number $u\in\mathbb{R}$, define *the first order difference operator with step size u in direction of the variable x_j* by

$$\Delta_{u,j}f(x)=f(x+u\mathbf{e}_j)-f(x),\quad j=1,\ldots,d.$$

By induction, the k-th order difference operator with step size u in direction of the variable x_j is defined as

$$\Delta_{u,j}^{k}f(x)=\Delta_{u,j}\Delta_{u,j}^{k-1}f(x)=\sum_{l=1}^{k}(-1)^{l+k}\binom{k}{l}\Delta_{ul,j}f(x). \tag{2.1}$$

Definition 1. *For given vectors $\vec{r}=(r_1,\ldots,r_d)\in[1,\infty]^d$ $\vec{\beta}=(\beta_1,\ldots,\beta_d)\in(0,\infty)^d$ and $\vec{L}=(L_1,\ldots,L_d)\in(0,\infty)^d$, we say that a function $f:\mathbb{R}^d\to\mathbb{R}^1$ belongs to the anisotropic Nikol'skii class $\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L}\big)$ if*

(i) $\|f\|_{r_j}\leq L_j$ *for all* $j=1,\ldots,d;$

(ii) *for every $j=1,\ldots,d$ there exists a natural number $k_j>\beta_j$ such that*

$$\left\|\Delta_{u,j}^{k_j}f\right\|_{r_j}\leq L_j|u|^{\beta_j},\quad\forall u\in\mathbb{R},\quad\forall j=1,\ldots,d.$$

Recall also that we introduced

$$\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L},M\big)=\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L}\big)\cap\big\{G:\mathbb{R}^d\to\mathbb{R}:\ \ \|G\|_\infty\leq M\big\},\ \ M>0.$$

Assumptions on the density g. Let $\mathfrak{I}^*$ denote the set of all subsets of $\{1,\ldots,d\}$. Set $\mathfrak{I}=\mathfrak{I}^*\cup\emptyset$ and for any $J\in\mathfrak{I}$ let $|J|$ denote the cardinality of J while $\{j_1<\cdots<j_{|J|}\}$ denotes its elements. For any $J\in\mathfrak{I}^*$ define the operator $\mathfrak{D}^J=\frac{\partial^{|J|}}{\partial t_{j_1}\cdots\partial t_{j_{|J|}}}$ and let $\mathfrak{D}^{\emptyset}$ denote the identity operator. For any $Q\in\mathbb{L}_1\big(\mathbb{R}^d\big)$ let $\check{Q}(t),t\in\mathbb{R}^d$, be the Fourier transform of Q.

Assumption 1. *$\mathfrak{D}^J\check{g}$ exists for any $J\in\mathfrak{I}^*$ and*

1) if $\alpha\in[0,1)$ then there exists $\mathfrak{d}_1>0$ such that

$$\sup_{J\in\mathfrak{I}^*}\big\|\mathfrak{D}^J\check{g}\big\|_\infty\leq\mathfrak{d}_1;$$

2) if $\alpha=1$ then there exists $\mathfrak{d}_2>0$ such that

$$\sup_{J\in\mathfrak{I}^*}\big\|\check{g}^{-1}\mathfrak{D}^J\check{g}\big\|_\infty\leq\mathfrak{d}_2.$$

Let us remark that Assumption 1 is very weak if $\alpha\in[0,1)$ and it is verified for many distributions, including multivariate centered Laplace and Gaussian ones. In the case $\alpha=1$ this assumption is much more restrictive. In particular, it does not hold for Gaussian law but it is still checked for the Laplace distribution.

Assumption 2. *If $\alpha = 1$ then there exist $\vec{\mu} = (\mu_1, \dots, \mu_d) \in (0, \infty)^d$ and $\Upsilon > 0$ such that*

$$\big|\check{g}(t)\big| \le \Upsilon \prod_{j=1}^{d} (1 + t_j^2)^{-\frac{\mu_j}{2}}, \quad \forall t = (t_1, \dots, t_d) \in \mathbb{R}^d.$$

The consideration of the $\mathbb{L}_p$-risk on $\mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}\big) \cap \mathbb{F}_g(R)$ requires to enforce Assumption 1 in the case $\alpha = 1$. For any $I, J \in \mathfrak{I}$ define $\mathfrak{D}^{I,J} = \mathfrak{D}^{I}\big(\mathfrak{D}^{J}\big)$ and note that obviously $\mathfrak{D}^{I,J} = \mathfrak{D}^{J,I}$.

Assumption 3. *$\mathfrak{D}^{I,J}\check{g}$ exists for any $I, J \in \mathfrak{I}$ and*

$$\textbf{(i)} \quad \sup_{I,J \in \mathfrak{I}} \big\|\mathfrak{D}^{I,J}\big(\check{g}\big)\big\|_1 =: \mathfrak{d}_3 < \infty;$$

$$\textbf{(ii)} \quad \sup_{J \in \mathfrak{I}^*} \int_{\mathbb{R}^d} g(z) \Big(\prod_{j \in J} z_j^2\Big) \mathrm{d}z < \infty.$$

Assumption 3 is not much restrictive and it is checked for multivariate centered Gaussian and Laplace distributions.

3. Main results.

To present our results in an unified way, let us define $\boldsymbol{\vec{\mu}}(\alpha) = \vec{\mu}$, $\alpha = 1$, $\boldsymbol{\vec{\mu}}(\alpha) = (0, \dots, 0)$, $\alpha \in [0, 1)$, and introduce the following quantities.

$$\frac{1}{\beta(\alpha)} = \sum_{j=1}^{d} \frac{2\boldsymbol{\mu}_j(\alpha) + 1}{\beta_j}, \quad \frac{1}{\omega(\alpha)} = \sum_{j=1}^{d} \frac{2\boldsymbol{\mu}_j(\alpha) + 1}{\beta_j r_j}, \quad L(\alpha) = \prod_{j=1}^{d} L_j^{\frac{2\boldsymbol{\mu}_j(\alpha)+1}{\beta_j}}.$$

Define for any $1 \le s \le \infty$ and $\alpha \in [0, 1]$

$$\varkappa_\alpha(s) = \omega(\alpha)(2 + 1/\beta(\alpha)) - s \qquad \tau(s) = 1 - 1/\omega(0) + 1/(s\beta(0)).$$

Convolution structure density model. Set $p^* = \big[\max_{l=1,\dots,d} r_l\big] \vee p$ and introduce

$$\varrho(\alpha) = \begin{cases} \frac{1-1/p}{1-1/\omega(\alpha)+1/\beta(\alpha)}, & \varkappa_\alpha(p) > p\omega(\alpha); \\ \frac{\beta(\alpha)}{2\beta(\alpha)+1}, & 0 < \varkappa_\alpha(p) \le p\omega(\alpha); \\ \frac{\tau(p)}{(2+1/\beta(\alpha))\tau(\infty)+1/[\omega(\alpha)\beta(0)]}, & \varkappa_\alpha(p) \le 0, \ \ \tau(p^*) > 0; \\ \frac{\omega(\alpha)(1-p^*/p)}{\varkappa_\alpha(p^*)}, & \varkappa_\alpha(p) \le 0, \ \ \tau(p^*) \le 0. \end{cases}$$

Here and later, we assume $0/0 = 0$, which implies in particular that $\frac{\omega(\alpha)(1-p^*/p)}{\varkappa_\alpha(p^*)} = 0$ if $p^* = p$ and $\varkappa_\alpha(p) = 0$. Also, the following observation will be useful in the sequel:

$$\varrho(\alpha) = \min\left[\frac{1 - \frac{1}{p}}{1 - \frac{1}{\omega(\alpha)} + \frac{1}{\beta(\alpha)}}, \frac{1}{2 + \frac{1}{\beta(\alpha)}}, \frac{\tau(p)}{\big(2 + \frac{1}{\beta(\alpha)}\big)\tau(\infty) + \frac{1}{\omega(\alpha)\beta(0)}}, \frac{\omega(\alpha)(1 - \frac{p^*}{p})}{\varkappa_\alpha(p^*)}\right]. \tag{3.1}$$

Put finally

$$\delta_n = \begin{cases} L(\alpha)n^{-1}, & \varkappa_\alpha(p) > 0; \\ L(\alpha)n^{-1}\ln(n), & \varkappa_\alpha(p) \le 0, \ \ \tau(p^*) \le 0; \\ [L(0)]^{-\frac{\varkappa_\alpha(p)}{\omega(\alpha)p\tau(p)}} L(\alpha)n^{-1}\ln(n), & \varkappa_\alpha(p) \le 0, \ \ \tau(p^*) > 0. \end{cases}$$

Theorem 1. *Let $L_0 > 0$ and $1 \le p < \infty$ be fixed. Then for any $\vec{\beta} \in (0,\infty)^d$, $\vec{r} \in [1,\infty]^d$, $\vec{L} \in [L_0,\infty)^d$, $R > 1$, $\vec{\mu} \in (0,\infty)^d$ and $g \in \mathfrak{P}(\mathbb{R}^d)$, satisfying Assumptions 1–3, there exists $c > 0$ independent of $\vec{L}$ such that*

$$\liminf_{n\to\infty} \inf_{\tilde{f}_n} \sup_{f\in\mathbb{N}_{\vec{r},d}(\vec{\beta},\vec{L})\cap\mathbb{F}_g(R)} \delta_n^{-\varrho(\alpha)} \mathcal{R}_p^{(n)}\big[\tilde{f}_n; f\big] \ge c,$$

where infimum is taken over all possible estimators. If additionally $\vec{\boldsymbol{\mu}}(\alpha) \in [0,1/2)^d$ in the case $\varkappa_\alpha(p) \le 0$, then

$$\liminf_{n\to\infty} \inf_{\tilde{f}_n} \sup_{f\in\mathbb{N}_{\vec{r},d}(\vec{\beta},\vec{L})\cap\mathfrak{P}(\mathbb{R}^d)} \delta_n^{-\varrho(\alpha)} \mathcal{R}_p^{(n)}\big[\tilde{f}_n; f\big] \ge c,$$

Following the terminology used in Goldenshluger and Lepski (2014), we will call the set of parameters satisfying $\varkappa_\alpha(p) > p\omega(\alpha)$ the tail zone, satisfying $0 < \varkappa_\alpha(p) \le p\omega(\alpha)$ the dense zone and satisfying $\varkappa_\alpha(p) \le 0$ the sparse zone. In its turn, the latter zone is divided in two sub-domains: the sparse zone 1 corresponding to $\tau(p^*) > 0$ and the sparse zone 2 corresponding to $\tau(p^*) \le 0$. The notion of dense and sparse zones appeared in the statistical literature in the mid-nineties in Donoho et al. (1996) and go up to the wavelet decomposition of the underlying density. The tail zone was discovered in Juditsky and Lambert–Lacroix (2004) and its existence is related to the function estimation on the whole space. In particular the consideration of compactly supported densities allows to eliminate this zone. The fact that the sparse zone is divided in two sub-domains was discovered in Goldenshluger and Lepski (2014) (bounded case) and in Lepski (2015) (unbounded case). This division can be informally viewed as some "degree of sparsity". It can be easily observed by analyzing the family of functions on which the lower bound established in the theorem is proved.

Some remarks are in order.

1^0. In view of the inclusion $\mathfrak{P}(\mathbb{R}^d) \subset \mathbb{F}_g(R)$ for any $g \in \mathfrak{P}(\mathbb{R}^d)$ and $R > 1$ we remark that the second assertion of the theorem is stronger than the first one. In particular, it provides with the lower bound for the minimax risk in the generalized deconvolution model (1.1). Note that the additional condition $\vec{\boldsymbol{\mu}}(\alpha) \in [0,1/2)^d$ is always satisfied if $\alpha \ne 1$ since $\vec{\boldsymbol{\mu}}(\alpha) = 0$. Moreover, for any $\alpha \in [0,1)$ the asymptotics of the of $\mathbb{L}_p$-risk found in Theorem 1 are independent of α. Also, it is worth mentioning that all our results in the case $\alpha \in (0,1)$ are proved under the very weak Assumption 1.

We would like to emphasize that the asymptotics of the $\mathbb{L}_p$-risk defined on $\mathbb{N}_{\vec{r},d}(\vec{\beta},\vec{L}) \cap \mathfrak{P}(\mathbb{R}^d)$ in the case $\vec{\boldsymbol{\mu}}(\alpha) \notin [0,1/2)^d$, $\alpha = 1$, on the sparse zone in the observation model (1.1) remains an open problem. However, as it will follow from Theorem 2 below the asymptotics found in Theorem 1 in the tail and the dense zone are correct on $\mathbb{N}_{\vec{r},d}(\vec{\beta},\vec{L}) \cap \mathfrak{P}(\mathbb{R}^d)$ as well, whatever the value of $\vec{\boldsymbol{\mu}}(\alpha)$.

2^0. The assertions of Theorem 1 in the case $\alpha \ne 0$ are completely new. In the case $\alpha = 0$, the lower bounds from Theorem 1 coincide with those found in Goldenshluger and Lepski (2014) when the tail or dense zone are considered. The result corresponding to the sparse zone even for $\alpha = 0$ was not known. Let us discuss this zone more in detail.

Case $\boldsymbol{\varkappa_\alpha(p) \le 0,\ \tau(p^*) \le 0,\ p^* = p}$. As it follows from Theorem 1 there is no uniformly consistent estimator under the $\mathbb{L}_p$-loss, $1 < p < \infty$, over $\mathbb{N}_{\vec{r},d}(\vec{\beta},\vec{L}) \cap \mathbb{F}_g(R)$ or over $\mathbb{N}_{\vec{r},d}(\vec{\beta},\vec{L}) \cap \mathfrak{P}(\mathbb{R}^d)$ if $\vec{\boldsymbol{\mu}}(\alpha) \in [0,1/2)^d$. We will see that the latter condition is necessary and sufficient for nonexistence of uniformly consistent estimators over these two functional classes.

Case $\boldsymbol{\varkappa_\alpha(p) \le 0,\ \tau(p^*) \le 0,\ p^* > p}$. It is interesting to note that this case does not appear in the dimension 1 or, more generally, when isotropic Nikol'skii classes are considered. Indeed, if $r_l = r$ for all $l = 1, \dots d$, then $p^* > p$ means $r > p$ which, in its turn, implies $\tau(p^*) = \tau(r) = 1 > 0$.

3^0. Note that if $p=1$, then necessarily $\varkappa_\alpha(p) > p\omega(\alpha)$ (the tail zone) and, therefore, a uniformly consistent estimator under $\mathbb{L}_1$-loss does not exist.

4^0. Let $\alpha = 1$ and let $p=2$ and $r_l = 2$ for any $l = 1, \dots, d$. It is easy to check that whatever the value of $\vec{\beta}$ and $\vec{\mu}$, the corresponding set of parameters belongs to the dense zone. The lower bound for the $\mathbb{L}_2$-risk in the deconvolution model was recently obtained in Comte and Lacour (2013) but under more restrictive assumption imposed on the density g. It is worth noting that in this case the asymptotics found in Theorem 1 is the minimax rate of convergence, see Comte and Lacour (2013).

Deconvolution density model. Bounded case. The problem that we address now is inspired by the following observation. Looking at (3.1) we remark that the elimination of any zone will improve the accuracy of estimation. A similar problem was investigated in the recent paper Goldenshluger and Lepski (2014), where in the direct case $\alpha = 0$ the authors proposed the tail dominance condition which allows to reduce and sometimes to eliminate the tail zone. In particular under this assumption the uniformly consistent estimation under $\mathbb{L}_1$-loss is possible.

Our goal is to impose in some sense minimal assumptions on the underlying function which would allow to eliminate the inconsistency zone $\varkappa_\alpha(p) \le 0,\ \tau(p^*) \le 0,\ p^* = p$. It turns out that in order to do this it suffices to add to the Nikol'skii class the uniform boundedness of the underlying function. Thus, our objective is to find a lower bound for minimax risk defined on $\mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}, M\big) \cap \mathfrak{P}\big(\mathbb{R}^d\big)$. Introduce

$$\rho(\alpha) \quad = \quad \begin{cases} \frac{1-1/p}{1-1/\omega(\alpha)+1/\beta(\alpha)}, & \varkappa_\alpha(p) > p\omega(\alpha); \\ \frac{\beta(\alpha)}{2\beta(\alpha)+1}, & 0 < \varkappa_\alpha(p) \le p\omega(\alpha); \\ \frac{\tau(p)}{(2+1/\beta(\alpha))\tau(\infty)+1/[\omega(\alpha)\beta(0)]}, & \varkappa_\alpha(p) \le 0,\ \ \tau(\infty) > 0; \\ \frac{\omega(\alpha)}{p}, & \varkappa_\alpha(p) \le 0,\ \tau(\infty) \le 0. \end{cases}$$

Similarly to (3.1)

$$\rho(\alpha) = \min\left[\frac{1-\frac{1}{p}}{1-\frac{1}{\omega(\alpha)}+\frac{1}{\beta(\alpha)}}, \frac{1}{2+\frac{1}{\beta(\alpha)}}, \frac{\tau(p)}{\big(2+\frac{1}{\beta(\alpha)}\big)\tau(\infty)+\frac{1}{\omega(\alpha)\beta(0)}}, \frac{\omega(\alpha)}{p}\right]. \tag{3.2}$$

Theorem 2. *Let $L_0 > 0$ and $1 \le p < \infty$ be fixed. Then for any $\vec{\beta} \in (0,\infty)^d$, $\vec{r} \in [1,\infty]^d$, $\vec{L} \in [L_0,\infty)^d$, $M > 0$, $\vec{\mu} \in (0,\infty)^d$ and $g \in \mathfrak{P}\big(\mathbb{R}^d\big)$, satisfying Assumptions 1 and 2 there exists $c > 0$ independent of $\vec{L}$ such that*

$$\liminf_{n\to\infty} \inf_{\tilde{f}_n} \sup_{f \in \mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L},M\big)\cap\mathfrak{P}\big(\mathbb{R}^d\big)} \delta_n^{-\rho(\alpha)} \mathcal{R}_p^{(n)}\big[\tilde{f}_n; f\big] \ge c,$$

where the infimum is taken over all possible estimators.

Remark 1. *The result of Theorem 2 in the case $\alpha \neq 0$ is completely new. When $\alpha = 0$, the lower bounds from Theorem 2 coincide with those found in Goldenshluger and Lepski (2014) when the tail, dense or sparse zone 1 are considered. As to the sparse zone 2, even if $\alpha = 0$, we improve in order the result obtained in Goldenshluger and Lepski (2014). Indeed, the asymptotics found in Goldenshluger and Lepski (2014) is proportional to $n^{-\rho(0)}$, which is smaller in order than $\delta_n^{\rho(0)}$. Moreover, we will see that the asymptotics given by $\delta_n^{\rho(\alpha)}$ yields the minimax rate of convergence on the whole sparse zone for any $\alpha \in [0,1]$.*

Comparing the assertions of Theorems 1 and 2 we conclude that the results of Theorem 2 concerning the tail and the dense zone imply those of Theorem 1. Moreover, their proofs do not require Assumption 3.

Let us also remark that the rate-index $\rho(\alpha)$ coincides with $\varrho(\alpha)$ if $p^* = \infty$, which is not a simple coincidence. Indeed, since $p < \infty$ then $p^* = \infty$ means that there exists $l \in \{1, \dots, d\}$ such that $r_l = \infty$. Hence, any function belonging to the corresponding Nikol'skii class is bounded by L_l in view of the definition of the class. Note, however, that the condition $p^* = \infty$ is much stronger than the belonging of the underlying function to $\mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}, M\big)$.

What is more, analyzing the proof of Theorem 1 in the case $\varkappa_\alpha(p) \le 0, \tau(p) \le 0, p^* = p$, we can conclude that

$$\liminf_{n\to\infty} \lim_{M\to\infty} \inf_{\tilde{f}_n} \sup_{f\in\mathbb{N}_{\vec{r},d}\left(\vec{\beta},\vec{L},M\right)\cap\mathfrak{P}\left(\mathbb{R}^d\right)} \mathcal{R}_p^{(n)}\big[\tilde{f}_n; f\big] > 0.$$

It shows that in some sense, the assumption $f \in \mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}, M\big)$ is the minimal one allowing to assert the existence of a uniformly consistent estimator on the whole sparse zone.

Bounds for the $\mathbb{L}_\infty$-risk. Introduce $\rho = \frac{\tau(\infty)}{(2+1/\beta(\alpha))\tau(\infty)+1/[\omega(\alpha)\beta(0)]}$.

Theorem 3. *Let $L_0 > 0$, $\vec{\beta} \in (0, \infty)^d$, $\vec{r} \in [1, \infty]^d$, $\vec{L} \in [L_0, \infty)^d$, $M > 0$, $\vec{\mu} \in (0, \infty)^d$ and $g \in \mathfrak{P}\big(\mathbb{R}^d\big)$, satisfying Assumptions 1 and 2, be fixed. Then,*

1) if $\tau(\infty) > 0$ there exists $c > 0$ independent of $\vec{L}$ such that

$$\liminf_{n\to\infty} \inf_{\tilde{f}_n} \sup_{f\in\mathbb{N}_{\vec{r},d}\left(\vec{\beta},\vec{L}\right)\cap\mathfrak{P}\left(\mathbb{R}^d\right)} \Big([L(0)]^{\frac{1}{\omega(\alpha)\tau(\infty)}} L(\alpha) n^{-1} \ln(n) \Big)^{-\rho} \mathcal{R}_\infty^{(n)}\big[\tilde{f}_n; f\big] \ge c,$$

where the infimum is taken over all possible estimators;

2) there is no uniformly consistent estimator if $\tau(\infty) \le 0$.

The assertions of the theorem are new if $\alpha \neq 0$. If $\alpha = 0$ the same bound was obtained in Goldenshluger and Lepski (2014) and the asymptotics found in the first assertion is the minimax rate of convergence, see Lepski (2013). In the univariate case $d = 1$, if $\alpha = 1$, and $r = \infty$ (Hölder class) the first assertion of the theorem was proved in Lounici and Nickl (2011).

Remark 2. *In view of the embedding theorem for Nikol'skii spaces, Nikol'skii (1977), Section 6.9, the condition $\tau(\infty) > 0$ guarantees that all the functions belonging to $\mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}\big)$ are uniformly bounded by a constant completely determined by $\vec{L}$. In view of the latter remark we can assert that Theorem 3 is a particular case of Theorem 2. However, although the proofs of these theorem have a lot of common elements, they are different. At least, the authors were unable to find a single proof for both theorems.*

4. Are the lower bounds sharp?

Since all the results presented in the previous section deal with lower bounds for the minimax risk only, a natural question is: *are the asymptotics that we found optimal?* In other words is it possible to construct an estimator whose accuracy coincides up to numerical constants with the asymptotics established in Theorems 1–3? As it was mentioned above, the development of rate-optimal estimators will be addressed in a separate paper. Here we would like to discuss some general ideas which, at a glance, will lead to the construction of an estimator which is adaptive over the scale of Nikol'skii classes. Moreover we consider some particular cases where the construction of an estimator attaining our lower bounds is straightforward.

Below we will discuss only the results of Theorems 1 and 2. An estimator attaining simultaneously the asymptotics proved in Theorem 3 was recently constructed in Rebelles (2015). The latter implies that our results concerning the estimation under sup-norm loss are sharp.

4.1. General conjecture.

For any $\vec{h}=(h_1,\dots,h_d)\in(0,\infty)^d$ set $V_{\vec{h}}=\prod_{j=1}^d h_j$ and introduce

$$K_{\vec{h}}(t)=V_{\vec{h}}^{-1}K\big(t_1/h_1,\dots,t_d/h_d\big),\ t\in\mathbb{R}^d,$$

where $K:\mathbb{R}^d\to\mathbb{R}$ is a kernel obeying some restrictions discussed later.

For any $\vec{h}\in(0,\infty)^d$ let $M\big(\cdot,\vec{h}\big)$ satisfy the operator equation

$$K_{\vec{h}}(y)=(1-\alpha)M\big(y,\vec{h}\big)+\alpha\int_{\mathbb{R}^d}g(t-y)M\big(t,\vec{h}\big)\mathrm{d}t,\quad y\in\mathbb{R}^d. \tag{4.1}$$

Introduce for any $\vec{h}\in(0,\infty)^d$ and $x\in\mathbb{R}^d$ the kernel estimator

$$\widehat{f}_{\vec{h}}(x)=n^{-1}\sum_{i=1}^n M\big(Z_i-x,\vec{h}\big). \tag{4.2}$$

Let $\mathcal{H}^d$ denote the diadic grid in $(0,\infty)^d$. We believe that for any given $x\in\mathbb{R}^d$ a data-driven selection rule from the family of kernel estimators $\big\{\widehat{f}_{\vec{h}}(x),\ \vec{h}\in\mathcal{H}^d\big\}$ will attain simultaneously

- ► the rates found in Theorems 1–2 on the whole sparse zone;
- ► the rates found in Theorems 1–2 up to a logarithmic factor on the tail and the dense zones,

under the following assumptions imposed on the density g.

1) If $\alpha=1$ then there exists $\vec{\mu}=(\mu_1,\dots,\mu_d)\in(0,\infty)^d$ and $\Upsilon_0>0$ such that

$$|\check{g}(t)|\geq\Upsilon_0\prod_{j=1}^d(1+t_j^2)^{-\frac{\mu_j}{2}},\quad\forall t=(t_1,\dots,t_d)\in\mathbb{R}^d. \tag{4.3}$$

Comparing (4.3) with Assumption 2 we can assert that they both are coherent. Indeed, we come to the following assumption

$$\Upsilon_0\prod_{j=1}^d(1+t_j^2)^{-\frac{\mu_j}{2}}\leq|\check{g}(t)|\leq\Upsilon\prod_{j=1}^d(1+t_j^2)^{-\frac{\mu_j}{2}},\quad\forall t\in\mathbb{R}^d.$$

This assumption is well-known in the literature, and corresponds to a *moderately ill-posed* statistical inverse problem. In particular, it is checked for the centered multivariate Laplace law.

2) If $\alpha\in(0,1)$ then there exists $\varepsilon>0$ such that

$$\big|1-\alpha+\alpha\check{g}(t)\big|\geq\varepsilon,\quad\forall t\in\mathbb{R}^d. \tag{4.4}$$

Note first that (4.4) is verified for many distributions. In particular, it holds with $\varepsilon=1-\alpha$ if $\check{g}$ is a real positive function. The latter is true, in particular, for centered multivariate Laplace and Gaussian laws and for any probability law obtained by an even number of convolutions of a symmetric distribution with itself. Moreover this condition always holds with $\varepsilon = 1-2\alpha$ if $\alpha < 1/2$.

The selection scheme that we have in mind is a generalization of two pointwise procedures. The first one is developed in Kerkyacharian et al. (2001) and the second one in Goldenshluger and Lepski (2014). In particular, the kernel K used in the operator equation (4.1) has to be chosen as it was suggested in these aforementioned papers.

4.2. Particular cases.

Let $\vec{r} = (p, \dots, p)$. It is easily seen that in this case $\varkappa_\alpha(p) = 2\beta(\alpha)p > 0$, $\omega(\alpha) = \beta(\alpha)p$ and, therefore, independently of $\vec{\beta}$ and α the case $p \geq 2$ corresponds to the dense zone and $p < 2$ describes the tail zone. Let $\widehat{f}_{\vec{h}}$, $\vec{h} \in (0,1)^d$, be the kernel estimator defined in (4.2). Introduce

$$b_{\vec{h}}(f,\cdot) = \int_{\mathbb{R}^d} K_{\vec{h}}(u-\cdot)f(u)\mathrm{d}u - f(\cdot); \qquad \xi_{\vec{h}}(f,\cdot) = n^{-1}\sum_{i=1}^{n}\Big[M\big(Z_i - x, \vec{h}\big) - \mathbb{E}_f M\big(Z_i - x, \vec{h}\big)\Big].$$

Using the relation (4.1) one easily gets

$$\mathcal{R}_p^{(n)}\big[\hat{f}_{\vec{h}}, f\big] \leq \big\|b_{\vec{h}}(f,\cdot)\big\|_p + \Big(\mathbb{E}_f\big\|\xi_{\vec{h}}(f,\cdot)\big\|_p^p\Big)^{\frac{1}{p}}.$$

The choice of K proposed in in Kerkyacharian et al. (2001) and in Goldenshluger and Lepski (2014) implies, see for instance Goldenshluger and Lepski (2011), for any $1 \leq p < \infty$

$$\sup_{f\in\mathbb{N}_{\vec{r},d}(\vec{\beta},L)} \big\|b_{\vec{h}}(f,\cdot)\big\|_p \leq C_1 \sum_{j=1}^{d} L_j h_j^{\beta_j}, \quad \forall \vec{h} \in (0,1)^d.$$

Here C_1 is a constant independent of $\vec{h}$. The restriction $\vec{r} = (p\dots,p)$ is crucial for the latter bound.

Suppose additionally that the Fourier transform $\check{K}$ of K satisfies

$$\int_{\mathbb{R}^d}\big|\check{K}(t)\big|\prod_{j=1}^{d}(1+t_j^2)^{\frac{\mu_j(\alpha)}{2}}\mathrm{d}t \leq \mathbf{k}_1, \qquad \int_{\mathbb{R}^d}\big|\check{K}(t)\big|^2\prod_{j=1}^{d}(1+t_j^2)^{\mu_j(\alpha)}\mathrm{d}t \leq \mathbf{k}_2^2,$$

for some constants $\mathbf{k}_1 > 0$ and $\mathbf{k}_2 > 0$. Taking into account that

$$\mathbb{E}_f\big\|\xi_{\vec{h}}(f,\cdot)\big\|_p^p = \int_{\mathbb{R}^d}\mathbb{E}_f\Big(\big|\xi_{\vec{h}}(f,x)\big|^p\Big)\mathrm{d}x$$

and applying the Rozentahl inequality to the sum of i.i.d centered random variables $\xi_{\vec{h}}(f,\cdot)$ we get for any $p \geq 2$

$$\sup_{f\in\mathbb{N}_{\vec{r},d}(\vec{\beta},L)}\Big(\mathbb{E}_f\big\|\xi_{\vec{h}}(f,\cdot)\big\|_p^p\Big)^{\frac{1}{p}} \leq C_2 n^{-\frac{1}{2}}\prod_{j=1}^{d} h_j^{-\frac{1}{2}-\mu_j(\alpha)}, \quad \forall \vec{h} \in (0,1)^d.$$

Here C_2 is a constant independent of $\vec{h}$ and n. Thus, we have for any $p \geq 2$

$$\sup_{f\in\mathbb{N}_{\vec{r},d}(\vec{\beta},L)} \mathcal{R}_p^{(n)}\big[\hat{f}_{\vec{h}}, f\big] \leq C_1\sum_{j=1}^{d} L_j h_j^{\beta_j} + C_2 n^{-\frac{1}{2}}\prod_{j=1}^{d} h_j^{-\frac{1}{2}-\mu_j(\alpha)}, \quad \forall \vec{h} \in (0,1)^d.$$

Minimizing the right hand side of this inequality w.r.t $\vec{h} \in (0,1)^d$, we come to the rate found in Theorem 1, which corresponds to the dense zone.

The situation is more delicate when $1 < p < 2$ (recall that there is no consistent estimator under the $\mathbb{L}_1$-loss). Indeed, applying Bahr-Esseen inequality, von Bahr and Esseen (1965), we obtain

$$\sup_{f \in \mathbb{N}_{\vec{r},d}(\vec{\beta},L)} \Big(\mathbb{E}_f \big\|\xi_{\vec{h}}(f,\cdot)\big\|_p^p\Big)^{\frac{1}{p}} \le C_2 n^{-1+\frac{1}{p}} \big\|M\big(\cdot,\vec{h}\big)\big\|_p, \quad \forall \vec{h} \in (0,1)^d.$$

Under slightly weaker assumptions than Assumptions 1 and 3 (**i**), the following estimate is available for any $1 < p < 2$. There exists C_3 such that

$$\big\|M\big(\cdot,\vec{h}\big)\big\|_p \le C_3 \prod_{j=1}^d h_j^{\frac{1}{p}-1-\mu_j(\alpha)}, \quad \forall \vec{h} \in (0,1)^d.$$

Thus, we have

$$\sup_{f \in \mathbb{N}_{\vec{r},d}(\vec{\beta},L)} \mathcal{R}_p^{(n)}\big[\hat{f}_{\vec{h}}, f\big] \le C_1 \sum_{j=1}^d L_j h_j^{\beta_j} + C_4 n^{-1+\frac{1}{p}} \prod_{j=1}^d h_j^{\frac{1}{p}-1-\mu_j(\alpha)}, \quad \forall \vec{h} \in (0,1)^d.$$

Minimizing the right hand side of this inequality w.r.t $\vec{h} \in (0,1)^d$, we come to the rate found in Theorem 1, which corresponds to the tail zone in the case $\alpha \in [0,1)$. Contrary to this, if $\alpha = 1$, the minimum obtained is larger in order than the asymptotics given in Theorem 1. We conjecture that the result of Theorem 1 is sharp. This implies, in particular, that in the pure deconvolution model the linear estimators are no more rate-optimal (minimax) whenever $p < 2$.

5. Proof of Theorems 1–2.

The proofs below are very long, technical and tricky and we break them in several parts.

In Section 5.2 we present three different constructions of finite sets of functions on which the announced lower bounds are established. All three constructions are used in the different parts of the proof of Theorem 1 while the proof of Theorem 2 uses the first construction only. Each construction contains several parameters to be chosen. We present the relationships between them allowing to assert that the corresponding set of functions belongs to the considered functional class.

Section 5.3 is devoted to the proof of a generic lower bound based on the constructions presented in Section 5.2. The results we obtain are explicitly expressed in terms of the aforementioned parameters and their proofs require to impose some additional restrictions on them. All conditions together with the obtained bounds are summarized in Section 5.3.3.

The proofs of Theorems 1 and 2 are given in Sections 5.4–5.6 and they consist in the optimal choice of the parameters satisfying the restrictions found in the previous sections.

5.1. Technical lemmas.

Set $T(x) = \pi^{-d} \prod_{j=1}^d \big(1 + x_j^2\big)^{-1}$, $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, and for any $N > 0$ and $\mathfrak{a} > 0$ define

$$\bar{f}_{0,N}(x) = (2N)^{-d} \int_{\mathbb{R}^d} T(y-x) 1_{[-N,N]^d}(y) \mathrm{d}y, \quad f_{0,N}(\cdot) = \mathfrak{a}^d \bar{f}_{0,N\mathfrak{a}}\big(\cdot\, \mathfrak{a}\big)$$

Lemma 1. *1) $f_{0,N}$ is a probability density for any value of N and $\mathfrak{a}$. Moreover, for any $\vec{\beta} \in (0,\infty)^d, \vec{r} \in (0,\infty]^d$ and $L_0 > 0$ there exists $\mathfrak{a} > 0$ such that*

$$f_{0,N}(\cdot) \in \mathbb{N}_{\vec{r},d}\big(\vec{\beta}, 2^{-1}\vec{L}_0\big), \quad \forall N > 0,$$

where $\vec{L}_0 = (L_0, \dots, L_0)$.

2) For any $M > 0$ one can find $N(M) > 0$ such that

$$f_{0,N}(\cdot) \in \mathbb{N}_{\vec{r},d}\big(\vec{\beta}, 2^{-1}\vec{L}_0, M\big), \quad \forall N \geq N(M).$$

3) Set $A = \mathfrak{a}\big[2\pi\{1+4\mathfrak{a}^2\}\big]^{-1}$. Then, for any $x \in \mathbb{R}^d$ and any $N \geq 2$

$$f_{0,N}(x) \geq A^d \prod_{j=1}^{d} \Big[N^{-1}1_{[-1-N,N+1]}(x_j) + \big(x_j^2 - N^2\big)^{-1}1_{\mathbb{R}\setminus[-1-N,N+1]}(x_j)\Big].$$

The first and second assertions of the lemma are obvious and the third one will be proven in the Appendix.

For any $g \in \mathfrak{P}\big(\mathbb{R}^d\big)$ let $N_g > 0$ be a given solution of the equation

$$\int_{[-N_g,N_g]^d} g(u)\mathrm{d}u = 2^{-1}.$$

For any $J \in \mathfrak{J}$ define

$$\Gamma_J = \big\{x \in \mathbb{R}^d : \ \ |x_j| > N + N_g + 1, \ j \in J, \ \ |x_j| \leq N + N_g + 1, \ j \in \bar{J}\big\}$$

where as usual $\bar{J} = \{1, \dots d\} \setminus J$ and later on the product over an empty set is supposed to be equal to 1.

Lemma 2. *For any $g \in \mathfrak{P}\big(\mathbb{R}^d\big)$ and any $N \geq 2$*

$$\big[f_{0,N} \star g\big](x) \geq B \sum_{J \in \mathfrak{J}} \Big(N^{|J|-d} \prod_{j \in J} x_j^{-2}\Big) 1_{\Gamma_J}(x), \quad \forall x \in \mathbb{R}^d,$$

where B depends on N_g and $\mathfrak{a}$ only.

5.2. Three general constructions of a finite set of functions.

Let $f^{(0)} = f_{0,N}$, where $f_{0,N}$ is constructed in Lemma 1 with the parameter $N = N_n > 8 \vee N(M/2)$ which will be chosen later. Recall that in view of the first and second assertions of Lemma 1

$$f^{(0)} \in \mathbb{N}_{\vec{r},d}\big(\vec{\beta}, 2^{-1}\vec{L}_0, 2^{-1}M\big) \cap \mathfrak{P}\big(\mathbb{R}^d\big). \tag{5.1}$$

Set $\mathfrak{n} = \big(\lfloor \max_{j=1,\dots,d} \beta_j \rfloor\big) \vee \big(\lfloor \max_{j=1,\dots,d} \mu_j \rfloor\big) + 3$ and let $\lambda : \mathbb{R}^1 \to \mathbb{R}^1, \ \lambda \in \mathbb{C}^{\mathfrak{n}}\big(\mathbb{R}^1\big), \ \lambda \not\equiv 0$ be a function possessing the following properties.

$$\mathrm{supp}(\lambda) \subseteq [-1,1], \qquad \boldsymbol{\lambda}_\infty = \sup_{k=0,\dots,\mathfrak{n}} \|\lambda^{(k)}\|_\infty < \infty. \tag{5.2}$$

Obviously, the second property is the consequence of the first one since $\lambda \in \mathbb{C}^{\mathfrak{n}}\big(\mathbb{R}^1\big)$.

For any $l=1,\dots,d$ let $1>\sigma_l=\sigma_l(n)\to 0$, $n\to\infty$, and $M_l\in\mathbb{N}^*$, $M_l\leq N(8\sigma_l)^{-1}+1$, be the sequences to be specified later. Define also

$$x_{j,l}=\big\{2(1-M_l)+4(j-1)\big\}\sigma_l,\quad j=1,\dots,M_l,\;\; l=1,\dots,d,$$

and let $\mathcal{M}=\{1,\dots,M_1\}\times\cdots\times\{1,\dots,M_d\}$. For any $\mathbf{m}=(m_1,\dots,m_d)\in\mathcal{M}$ introduce

$$\Lambda_{\mathbf{m}}(x)=\prod_{l=1}^{d}\lambda\left(\frac{x_l-x_{m_l,l}}{\sigma_l}\right),\qquad x\in\mathbb{R}^d,$$
$$\Pi_{\mathbf{m}}=\big[x_{m_1,1}-2\sigma_1,x_{m_1,1}+2\sigma_1\big]\times\cdots\times\big[x_{m_d,d}-2\sigma_d,x_{m_d,d}+2\sigma_d\big]\subset\mathbb{R}^d.$$

Note that in view of the first condition in (5.2), if $\dot{U}$ denotes the interior of a set U:

$$\operatorname{supp}\big(\Lambda_{\mathbf{m}}\big)\subset\Pi_{\mathbf{m}},\quad\forall m\in\mathcal{M},\tag{5.3}$$
$$\dot{\Pi}_{\mathbf{m}}\cap\dot{\Pi}_{\mathbf{j}}=\emptyset,\quad\forall\mathbf{m},\mathbf{j}\in\mathcal{M}:\;\;\mathbf{m}\neq\mathbf{j}.\tag{5.4}$$

For $\mathbf{m}\in\mathcal{M}$ define

$$\pi(\mathbf{m})=\sum_{j=1}^{d-1}(m_j-1)\bigg(\prod_{l=j+1}^{d}M_l\bigg)+m_d.$$

It is easily checked that π defines an enumeration of the set $\mathcal{M}$, and $\pi:\mathcal{M}\to\{1,2\dots,|\mathcal{M}|\}$ is a bijection. For any $w\in\{0,1\}^{|\mathcal{M}|}$ define

$$F_w(x)=\mathfrak{A}\sum_{\mathbf{m}\in\mathcal{M}}w_{\pi(\mathbf{m})}\Lambda_{\mathbf{m}}(x),\qquad x\in\mathbb{R}^d,$$

where w_s, $s=1,\dots,|\mathcal{M}|$ are the coordinates of w, and $\mathfrak{A}$ is a parameter to be specified.

Let W be a subset of $\{0,1\}^{|\mathcal{M}|}$. We remark that the construction of the set of functions $\big\{F_w,w\in\{0,1\}^{|\mathcal{M}|}\big\}$ almost coincides with the construction proposed in Goldenshluger and Lepski (2014), in the proof of Theorem 3. Thus, completely repeating the computations done in the latter paper we can assert that the assumption

$$\mathfrak{A}\sigma_l^{-\beta_l}\bigg(S_W\prod_{j=1}^{d}\sigma_j\bigg)^{1/r_l}\leq(2C_1)^{-1}L_l,\quad\forall l=1,\dots,d,\tag{5.5}$$

guarantees that $F_w\in\mathbb{N}_{\vec{r},d}(\vec{\beta},2^{-1}\vec{L})$ for any $w\in W$.

Here $S_W:=\sup_{w\in W}|\{s:\;w_s\neq 0\}|$ and C_1 is a universal constant completely determined by the function λ. It is important to mention that the proof of (5.5) uses only the condition (5.2), which is the same as in Goldenshluger and Lepski (2014).

5.2.1. First construction.

Suppose additionally that

$$\int_{\mathbb{R}^1}y^k\lambda(y)\mathrm{d}y=0,\;\forall k=0,\dots,\mathfrak{n}.\tag{5.6}$$

It is easy to see that the set of functions satisfying (5.2) and (5.6) is not empty. One of the possible constructions of such λ consists in the following. Define

$$\lambda(t)=\sum_{s=1}^{\mathfrak{n}+2}a_s\big(1-t^2\big)^{s+\mathfrak{n}}1_{[-1,1]}(t),\quad t\in\mathbb{R}^1.$$

It is obvious that $\lambda\in\mathbb{C}^{\mathfrak{n}}\big(\mathbb{R}^1\big)$ and it satisfies the two conditions in (5.2), whatever the values of $a_1,\ldots,a_{\mathfrak{n}+2}$. Condition (5.6) is reduced to

$$\sum_{s=1}^{\mathfrak{n}+2}a_sb_s(k)=0,\quad\forall k=0,\ldots,\mathfrak{n},$$

where we have put $b_s(k)=\int_{-1}^{1}\big(1-t^2\big)^{s+\mathfrak{n}}t^k\mathrm{d}t$.

The latter system of linear equations always has a solution $\big(a_1,\ldots,a_{\mathfrak{n}+2}\big)\neq(0,\ldots,0)$ since the number of equations $(\mathfrak{n}+1)$ is less than the number of variables $(\mathfrak{n}+2)$. It implies in particular that $\lambda\not\equiv0$ since λ is a polynomial on $[-1,1]$.

It follows from (5.2), (5.3) and (5.4) that

$$\big\|F_w\big\|_\infty\leq\mathfrak{A}\boldsymbol{\lambda}_\infty^d,\quad\forall w\in\{0,1\}^{|\mathcal{M}|},\tag{5.7}$$

and (5.6) with $k=0$ implies that

$$\int_{\mathbb{R}^d}F_w(x)\mathrm{d}x=0,\quad\forall w\in\{0,1\}^{|\mathcal{M}|}.\tag{5.8}$$

Define for any $w\in\{0,1\}^{|\mathcal{M}|}$

$$f_w(x)=f^{(0)}(x)+F_w(x),\qquad x\in\mathbb{R}^d.$$

and remind that $f^{(0)}\in\mathfrak{P}\big(\mathbb{R}^d\big)$. It yields, first, together with (5.8) for any $w\in\{0,1\}^{|\mathcal{M}|}$

$$\int_{\mathbb{R}^d}f_w(x)\mathrm{d}x=1.\tag{5.9}$$

Next, under (5.5) $f_w\in\mathbb{N}_{\vec{r},d}(\vec{\beta},\vec{L})$ for any $w\in W$ in view of (5.1) since $\min_{j=1,\ldots,d}L_j\geq L_0$.

At last, if we impose the restriction $M_l\leq N(8\sigma_l)^{-1}+\frac{1}{2}$, then by construction of F_w, for any $w\in\{0,1\}^{|\mathcal{M}|}$

$$F_w(x)=0,\qquad\forall x\notin[-N/4,N/4]^d.$$

This yields $f_w(x)=f^{(0)}(x)\geq0$ for all $x\notin[-N/4,N/4]^d$.

On the other hand, in view of the third assertion of Lemma 1

$$f^{(0)}(x)\geq AN^{-d},\qquad\forall x\in[-N/4,N/4]^d.$$

Therefore, if we require, putting $C_2=2^{-1}A\boldsymbol{\lambda}_\infty^{-d}$,

$$\mathfrak{A}\leq C_2N^{-d},\tag{5.10}$$

this will imply together with (5.7) for any $x \in [-N/4, N/4]^d$

$$f_w(x) \geq f^{(0)}(x) - \|F_w\|_\infty \geq 2^{-1} f^{(0)}(x) + 2^{-1} A N^{-d} - \|F_w\|_\infty \geq 2^{-1} f^{(0)}(x).$$

Thus, we have finally

$$f_w(x) \geq 2^{-1} f^{(0)}(x) > 0, \quad \forall x \in \mathbb{R}^d, \quad \forall w \in \{0,1\}^{|\mathcal{M}|}, \tag{5.11}$$

and we conclude that under (5.10) $f_w \geq 0$ for any $w \in \{0,1\}^{|\mathcal{M}|}$.

Thus, we assert in view of (5.9) and (5.11) that under (5.10) f_w is a probability density for any $w \in \{0,1\}^{|\mathcal{M}|}$. Thus, we can assert that under restrictions (5.5) and (5.10) $\{f^{(0)}, f_w, w \in W\}$ is a finite set of probability densities from $\mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L})$.

It remains to note that for any N providing $C_2 N^{-d} \boldsymbol{\lambda}_\infty^d \leq 2^{-1} M$ we guarantee $\|F_w\|_\infty \leq 2^{-1} M$ for any $w \in \{0,1\}^{|\mathcal{M}|}$ in view of (5.7) and (5.10). Thus, we conclude that under (5.5) and (5.10)

$$\big\{f^{(0)}, f_w, w \in W\big\} \subset \mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}, M\big) \cap \mathfrak{P}\big(\mathbb{R}^d\big). \tag{5.12}$$

Here we have also taken into account (5.1).

5.2.2. Second construction.

Instead of (5.6) suppose now that

$$\lambda(y) \geq 0, \ \forall y \in \mathbb{R}. \tag{5.13}$$

Introduce for any $w \in \{0,1\}^{|\mathcal{M}|}$

$$f_w(x) = \bigg(1 - \|\lambda\|_1^d \mathfrak{A} S(w) \prod_{l=1}^d \sigma_l\bigg) f^{(0)}(x) + F_w(x), \qquad x \in \mathbb{R}^d,$$

where we have put $S(w) = |\{s : \ w_s \neq 0\}|$. Noting that

$$\int_{\mathbb{R}^d} F_w(x) \mathrm{d}x = \|\lambda\|_1^d \mathfrak{A} S(w) \prod_{l=1}^d \sigma_l$$

in view of (5.3) and (5.4), we obtain using (5.1)

$$\int_{\mathbb{R}^d} f_w(x) \mathrm{d}x = 1, \quad \forall w \in \{0,1\}^{|\mathcal{M}|}. \tag{5.14}$$

If we require, putting $C_3 = 2^{-1} \|\lambda\|_1^{-d}$, that

$$\mathfrak{A} S_W \prod_{l=1}^d \sigma_l \leq C_3 \tag{5.15}$$

we obtain that $f_w \geq 0$ for any $w \in W$, which together with (5.14) yields $\big\{f^{(0)}, f_w, w \in W\big\} \subset \mathfrak{P}\big(\mathbb{R}^d\big)$. Here we have used that $F_w \geq 0$ for any $w \in \{0,1\}^{|\mathcal{M}|}$ in view of (5.13), and that $S_W = \sup_{w \in W} S(w)$.

Note also, that under (5.5) and (5.15) $\big\{f^{(0)}, f_w, w \in W\big\} \subset \mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}\big)$ in view of (5.1). Thus, we can conclude that if (5.5) and (5.15) are fulfilled

$$\big\{f^{(0)}, f_w, w \in W\big\} \subset \mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}\big) \cap \mathfrak{P}\big(\mathbb{R}^d\big). \tag{5.16}$$

Note that $S_W \prod_{l=1}^d \sigma_l \leq C N^d$, where C is a universal constant, and therefore, the restriction (5.15) is weaker in general than the condition (5.10). Thus, the latter restriction leads to the inclusion (5.12) which is stronger than the inclusion (5.16) obtained under (5.15).

5.2.3. Third construction.

This construction will be used only if $\alpha = 1$. We will assume that (5.6) holds and use the first construction, i.e. for any $w \in \{0,1\}^{|\mathcal{M}|}$

$$f_w(x) = f^{(0)}(x) + F_w(x), \qquad x \in \mathbb{R}^d.$$

We already proved that $\{f^{(0)}, f_w, w \in W\} \subset \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L})$ if (5.5) holds. Thus, it remains to find conditions guaranteeing $\{f^{(0)}, f_w, w \in W\} \subset \mathbb{F}_g(R)$. Obviously $f^{(0)} \in \mathbb{B}_{1,d}(R)$, so we need to find conditions assuring that (remember that $\alpha = 1$):

$$f_w \in \mathbb{B}_{1,d}(R) \tag{5.17}$$

$$\{[f^{(0)} \star g], [f_w \star g], w \in W\} \subset \mathfrak{P}(\mathbb{R}^d). \tag{5.18}$$

We start with finding condition guaranteeing the verification (5.17). We have:

$$\int_{\mathbb{R}^d} |f_w(x)|\mathrm{d}x \le \int_{\mathbb{R}^d} f^{(0)}(x)\mathrm{d}x + \int_{\mathbb{R}^d} |F_w(x)|\mathrm{d}x \le 1 + \|\lambda\|_1^d \mathfrak{A} S_W \prod_{l=1}^d \sigma_l.$$

Putting $C_{0,3} = (R-1)8^d\|\lambda\|_1^{-d}$ and imposing

$$\mathfrak{A} N^d S_W \prod_{l=1}^d \sigma_l \le C_{0,3} \tag{5.19}$$

we can assert that (5.17) holds (remember that $N \ge 8$).

Now, let us find conditions guaranteeing (5.18). Since $f^{(0)}, g \in \mathfrak{P}(\mathbb{R}^d)$ we have in view of (5.8)

$$\int_{\mathbb{R}^d} [f^{(0)} \star g](x)\mathrm{d}x = 1, \qquad \int_{\mathbb{R}^d} [f_w \star g](x)\mathrm{d}x = 1, \quad \forall w \in \{0,1\}^{|\mathcal{M}|}. \tag{5.20}$$

Moreover, $f^{(0)}, g \in \mathfrak{P}(\mathbb{R}^d)$ implies obviously $[f^{(0)} \star g] \ge 0$ and, therefore,

$$[f^{(0)} \star g] \in \mathfrak{P}(\mathbb{R}^d). \tag{5.21}$$

Thus, it remains to find conditions under which

$$[f_w \star g] \ge 0, \quad \forall w \in W. \tag{5.22}$$

1^0. We have for any $w \in W$ in view of the Young inequality

$$\|F_w \star g\|_\infty \le \|g\|_\infty \|F_w\|_1 \le \|g\|_\infty \|\lambda\|_1^d \mathfrak{A} S_W \prod_{l=1}^d \sigma_l.$$

In view of Lemma 2 $[f^0 \star g](x) \ge BN^{-d}$ for any $x \in \Gamma_\emptyset$ and, therefore, if we require that

$$\mathfrak{A} N^d S_W \prod_{l=1}^d \sigma_l \le C_{1,3} \tag{5.23}$$

holds with $C_{1,3} = (2\|g\|_\infty \|\lambda\|_1^d)^{-1} B$ we can assert that

$$[f_w \star g](x) \geq \big[f^{(0)} \star g\big](x) - \big\|F_w \star g\big\|_\infty \geq 2^{-1}\big[f^{(0)} \star g\big](x), \quad \forall x \in \Gamma_\emptyset. \tag{5.24}$$

2^0. Set $\mathbf{\Lambda}(x) = \prod_{l=1}^d \lambda(x_l/\sigma_l)$ and let $\check{\mathbf{\Lambda}}$ denote the Fourier transform of $\mathbf{\Lambda}$. Note that in view of Assumption 3 and the conditions (5.2) imposed on λ

$$\mathfrak{D}^{J,J}\big[\check{g}(t)\check{\mathbf{\Lambda}}(t)\big] = \sum_{\mathcal{I},\mathcal{J} \subseteq J \cup \emptyset} \mathfrak{D}^{\mathcal{I},\mathcal{J}}\big(\check{\mathbf{\Lambda}}(t)\big)\mathfrak{D}^{J\setminus\mathcal{I},J\setminus\mathcal{J}}\big(\check{g}(t)\big). \tag{5.25}$$

Moreover, $\check{\mathbf{\Lambda}}(t) = \prod_{l=1}^d \sigma_l \check{\lambda}\big(t_l\sigma_l\big)$ and, therefore for any $\mathcal{I},\mathcal{J} \subseteq J \cup \emptyset$

$$\mathfrak{D}^{\mathcal{I},\mathcal{J}}\big(\check{\mathbf{\Lambda}}(t)\big) = \prod_{l \in \mathcal{I}\cap\mathcal{J}} \sigma_l^3 \check{\lambda}''\big(t_l\sigma_l\big) \prod_{l \in \{\mathcal{I}\cup\mathcal{J}\}\setminus\{\mathcal{I}\cap\mathcal{J}\}} \sigma_l^2 \check{\lambda}'\big(t_l\sigma_l\big) \prod_{l \in J\setminus\{\mathcal{I}\cup\mathcal{J}\}} \sigma_l \check{\lambda}\big(t_l\sigma_l\big),$$

where $\check{\lambda}'$ and $\check{\lambda}''$ denote the first and the second derivative of $\check{\lambda}$ respectively.

Hence, we have, putting $\tilde{\lambda} = |\check{\lambda}| \vee |\check{\lambda}'| \vee |\check{\lambda}''|$ and taking into account that $\sigma_l \leq 1, l = 1, \ldots, d$,

$$\big|\mathfrak{D}^{\mathcal{I},\mathcal{J}}\big(\check{\mathbf{\Lambda}}(t)\big)\big| \leq \prod_{l=1}^d \sigma_l \tilde{\lambda}\big(t_l\sigma_l\big) \leq \|\tilde{\lambda}\|_\infty^d \prod_{l=1}^d \sigma_l, \quad \forall \mathcal{I},\mathcal{J} \subseteq J \cup \emptyset, \quad \forall t \in \mathbb{R}^d. \tag{5.26}$$

3^0. For any $J \in \mathfrak{J}$ set $Q_J(y) = \big(\prod_{j\in J} y_j^2\big)\big[\mathbf{\Lambda} \star g\big](y), y \in \mathbb{R}^d$ and note that $Q_J \in \mathbb{L}_1\big(\mathbb{R}^d\big)$. Indeed,

$$\int_{\mathbb{R}^d} |Q(y)|\mathrm{d}y \leq \Big(\prod_{l=1}^d \sigma_l\Big)\|\lambda\|_1^{d-|J|} \int_{\mathbb{R}^d} g(z) \prod_{j \in J}\bigg[\int_{\mathbb{R}} |\lambda(u)|\big(z - u\sigma_j\big)^2 \mathrm{d}u\bigg]\mathrm{d}z < \infty$$

in view of (5.2) and Assumption 3 (**ii**).

Let $\check{Q}_J$ denote the Fourier transform of Q_J which is well-defined since $Q_J \in \mathbb{L}_1\big(\mathbb{R}^d\big)$. Then, we have for any $t = (t_1, \ldots, t_d) \in \mathbb{R}^d$ in view of (5.25)

$$\check{Q}_J(t) = (-1)^{|J|}\mathfrak{D}^{J,J}\big[\check{g}(t)\check{\mathbf{\Lambda}}(t)\big] = (-1)^{|J|} \sum_{\mathcal{I},\mathcal{J} \subseteq J \cup \emptyset} \mathfrak{D}^{\mathcal{I},\mathcal{J}}\big(\check{\mathbf{\Lambda}}(t)\big)\mathfrak{D}^{J\setminus\mathcal{I},J\setminus\mathcal{J}}\big(\check{g}(t)\big).$$

It yields, together with (5.26) and Assumption 3 (**i**) for any $J \in \mathfrak{J}$

$$\begin{aligned} \big\|Q_J\big\|_\infty &\leq (2\pi)^{-d}\big\|\check{Q}_J\big\|_1 \leq (2\pi)^{-d} \sum_{\mathcal{I},\mathcal{J} \subseteq J \cup \emptyset} \big\|\mathfrak{D}^{\mathcal{I},\mathcal{J}}\big(\check{\mathbf{\Lambda}}\big)\big\|_\infty \big\|\mathfrak{D}^{J\setminus\mathcal{I},J\setminus\mathcal{J}}\big(\check{g}\big)\big\|_1 \\ &\leq \mathfrak{d}_3(2\pi)^{-d}\big(4\big\|\tilde{\lambda}\big\|_\infty\big)^d \prod_{l=1}^d \sigma_l. \end{aligned} \tag{5.27}$$

4^0. Let $J \in \mathfrak{J}^*$ be fixed and set $\widetilde{F}_{w,J}(x) = [F_w(x) \star g](x)\prod_{j\in J} x_j^2$.
We obviously have in view of the triangle inequality

$$\sup_{x \in \Gamma_J}\big|\widetilde{F}_{w,J}(x)\big| \leq \mathfrak{A} \sum_{\mathbf{m}\in\mathcal{M}} w_{\pi(\mathbf{m})} \sup_{x\in\Gamma_J}\big|\widetilde{\Lambda}_{\mathbf{m},J}(x)\big|,$$

where we have put $\widetilde{\Lambda}_{\mathbf{m},J}(x) = [\Lambda_{\mathbf{m}} \star g](x)\prod_{j\in J} x_j^2$.

Note that the definition of Γ_J implies for any $x = (x_1, \ldots, x_d) \in \Gamma_J$

$$x_l^2 \le 4(x_l - x_{m_l,l})^2, \;\; \forall l \in J, \; \forall \mathbf{m} \in \mathcal{M}. \tag{5.28}$$

Here we have also used that the definition of $x_{s,l}, s = 1, \ldots, M_l, l = 1, \ldots, d$, together with the restriction $M_l \le N(8\sigma_l)^{-1} + \frac{1}{2}$ imply that $|x_{m_l,l}| \le N/4$ for any $l = 1, \ldots, d$ and $\mathbf{m} \in \mathcal{M}$.

It yields for any $\mathbf{m} \in \mathcal{M}$ in view of the definition of $\Lambda_{\mathbf{m}}$

$$\sup_{x\in\Gamma_J} \big|\widetilde{\Lambda}_{\mathbf{m},J}\big| \le 4^{|J|} \sup_{x\in\Gamma_J} \big|\Lambda_{\mathbf{m}}(x)\big| \prod_{j\in J}(x_j - x_{m_l,l})^2 \le 4^d \sup_{y\in\mathbb{R}^d} \big|[\mathbf{\Lambda} \star g](y)\big| \prod_{j\in J} y_j^2 =: 4^d \|Q_J\|_\infty,$$

and, therefore, we deduce from (5.27) for any $w \in W$

$$\sup_{x\in\Gamma_J} \big|\widetilde{F}_{w,J}\big| \le \mathfrak{d}_3(2\pi)^{-d}\big(16\|\tilde{\lambda}\|_\infty\big)^d \mathfrak{A} S_W \prod_{l=1}^d \sigma_l.$$

We get finally

$$\big|F_w(x)\big| \le \Big(\prod_{j\in J} x_j^{-2}\Big)\mathfrak{d}_3(2\pi)^{-d}\big(16\|\tilde{\lambda}\|_\infty\big)^d \mathfrak{A} S_W \prod_{l=1}^d \sigma_l, \quad \forall x \in \Gamma_J.$$

In view of Lemma 2, $\big[f^0 \star g\big](x) \ge B\Big(N^{|J|-d}\prod_{j\in J} x_j^{-2}\Big)$, for any $x \in \Gamma_J$ and, therefore, if we require that

$$\mathfrak{A} N^d S_W \prod_{l=1}^d \sigma_l \le C_{2,3} \tag{5.29}$$

holds with $C_{2,3} = 2^{-1} B \mathfrak{d}_3^{-1}(2\pi)^d\big(16\|\tilde{\lambda}\|_\infty\big)^{-d}$ we can assert that for any $J \in \mathfrak{J}^*$

$$[f_w \star g](x) \ge \big[f^{(0)} \star g\big](x) - \big|[F_w \star g](x)\big| \ge 2^{-1}\big[f^{(0)} \star g\big](x), \quad \forall x \in \Gamma_J. \tag{5.30}$$

Since $\Gamma_J, J \in \mathfrak{J}$, is a partition of $\mathbb{R}^d$, (5.24) and (5.30) imply that if $\mathfrak{A} N^d S_W \prod_{l=1}^d \sigma_l \le C_{1,3} \wedge C_{2,3}$

$$[f_w \star g](x) \ge 2^{-1}\big[f^{(0)} \star g\big](x) \ge 0, \quad \forall x \in \mathbb{R}^d, \tag{5.31}$$

Set $C_3' = C_{0,3} \wedge C_{1,3} \wedge C_{2,3}$ and assume

$$\mathfrak{A} N^d S_W \prod_{l=1}^d \sigma_l \le C_3'. \tag{5.32}$$

It remains to note that (5.31) implies (5.22), which together with (5.19), (5.20) and (5.21) allows us to conclude that

$$\big\{f^{(0)}, f_w, w \in W\big\} \subset \mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}\big) \cap \mathbb{F}_g(R), \tag{5.33}$$

if the restrictions (5.5) and (5.32) are fulfilled.

5.3. Generic lower bound.

As before the notation $\mathbf{C_1}, \mathbf{C_2}, \ldots,$ is used for the constants independent of $\vec{L}$.

Let $\mathbf{P}$ denote the probability law on $\{0,1\}^{|\mathcal{M}|}$ such that

$$\mathbf{P}(w) = \varsigma^{\sum_{s=1}^{|\mathcal{M}|} w_s}(1-\varsigma)^{|\mathcal{M}|-\sum_{s=1}^{|\mathcal{M}|} w_s}, \quad w \in \{0,1\}^{|\mathcal{M}|},$$

where $\varsigma \in (0, 1/2]$ will be chosen later. Denote by $\mathbf{E}$ the mathematical expectation with respect to $\mathbf{P}$. Choose also

$$W = \left\{ w \in \{0,1\}^{|\mathcal{M}|} : \sum_{s=1}^{|\mathcal{M}|} w_s \leq 2\varsigma|\mathcal{M}| \right\}.$$

and note that this choice of W implies $S_W \leq 2\varsigma\mathcal{M}$. Hence, (5.5) will be fulfilled if

$$\mathfrak{A}\sigma_l^{-\beta_l}\left(2\varsigma|\mathcal{M}|\prod_{j=1}^{d}\sigma_j\right)^{1/r_l} \leq (2C_1)^{-1}L_l, \quad \forall l = 1,\ldots,d,$$

and (5.15) and (5.32) will be held if

$$\mathfrak{A}\varsigma|\mathcal{M}|\prod_{l=1}^{d}\sigma_l \leq 2^{-1}C_3, \qquad \mathfrak{A}\varsigma|\mathcal{M}|N^d\prod_{l=1}^{d}\sigma_l \leq 2^{-1}C_3'.$$

Choose $M_l = N(8\sigma_l)^{-1}, l = 1,\ldots,d$, assuming without loss of generality that M_l is an integer. It yields

$$|\mathcal{M}| = (N/8)^d\left(\prod_{l=1}^{d}\sigma_l\right)^{-1}. \tag{5.34}$$

Then, (5.5) and (5.32) are reduced respectively to

$$\begin{aligned} \mathfrak{A}\sigma_l^{-\beta_l}\left(\varsigma N^d\right)^{1/r_l} &\leq \mathbf{C_1}^{-1}L_l, \quad \forall l = 1,\ldots,d; &(5.35)\\ \mathfrak{A}N^{2d}\varsigma &\leq \mathbf{C_3}. &(5.36)\end{aligned}$$

We note also that (5.32) implies (5.15) if we replace C_3 and C_3' by $C_3 \wedge C_3'$ since $N > 1$. Hence, we conclude that (5.36) implies (5.15).

Moreover, let us suppose that $\varsigma \geq 4|\mathcal{M}|^{-1}$ and, therefore, necessarily

$$N^d\varsigma \geq 2^{3d+2}\prod_{l=1}^{d}\sigma_l. \tag{5.37}$$

Note also that since π is bijection the following inclusion holds for any $\mathbf{j} \in \mathcal{M}$

$$W \supset \left\{ w \in \{0,1\}^{|\mathcal{M}|} : \sum_{\mathbf{m}\in\mathcal{M},\mathbf{m}\neq\mathbf{j}} w_{\pi(\mathbf{m})} \leq 2\varsigma|\mathcal{M}| - 1 \right\} =: W_{\mathbf{j}}. \tag{5.38}$$

5.3.1. Generic lower bound via the first and the third constructions.

Let Σ be either $\mathbb{N}_{\vec{r},d}(\vec{\beta},\vec{L},M)\cap\mathfrak{P}(\mathbb{R}^d)$ or $\mathbb{N}_{\vec{r},d}(\vec{\beta},\vec{L})\cap\mathbb{F}_g(R)$. We have under (5.35) and either (5.10) or (5.36) in view of either (5.12) or (5.33) for any estimator $\hat{f}$

$$R(\hat{f}) := \sup_{f\in\Sigma}\mathbb{E}_f\|\hat{f}-f\|_p^p \geq \mathbf{E}1_W\mathbb{E}_{f_w}\|\hat{f}-f_w\|_p^p.$$

1^0. Denote $\tilde{f}=\hat{f}-f^{(0)}$ and remark that

$$\|\hat{f}-f_w\|_p^p \geq \sum_{\mathbf{j}\in\mathcal{M}}\int_{\Pi_{\mathbf{j}}}\big|\tilde{f}(x)-F_w\big|^p\mathrm{d}x = \sum_{\mathbf{j}\in\mathcal{M}}\int_{\Pi_{\mathbf{j}}}\big|\tilde{f}(x)-\mathfrak{A}w_{\pi(\mathbf{j})}\Lambda_{\mathbf{j}}(x)\big|^p\mathrm{d}x,$$

in view of (5.3) and (5.4). Thus, we get

$$R(\hat{f}) \geq \sum_{\mathbf{j}\in\mathcal{M}}\int_{\Pi_{\mathbf{j}}}\mathbf{E}1_{W_{\mathbf{j}}}\Big(\mathbb{E}_{f_w}\big|\tilde{f}(x)-\mathfrak{A}w_{\pi(\mathbf{j})}\Lambda_{\mathbf{j}}(x)\big|^p\Big)\mathrm{d}x. \tag{5.39}$$

Set for any $\mathbf{j}\in\mathcal{M}$

$$f^{\mathbf{j}}_{w,0}(\cdot)=f^{(0)}+\mathfrak{A}\sum_{\mathbf{m}\in\mathcal{M},\mathbf{m}\neq\mathbf{j}}w_{\pi(\mathbf{m})}\Lambda_{\mathbf{m}}(\cdot), \qquad f^{\mathbf{j}}_{w,1}(\cdot)=f^{\mathbf{j}}_{w,0}(\cdot)+\mathfrak{A}\Lambda_{\mathbf{j}}(\cdot)$$

and define

$$\mathfrak{Z}_{\mathbf{j}}(w) := \frac{\mathrm{d}\mathbb{P}_{f^{\mathbf{j}}_{w,1}}}{\mathrm{d}\mathbb{P}_{f^{\mathbf{j}}_{w,0}}}\big(Z^{(n)}\big) = \prod_{k=1}^n\frac{(1-\alpha)f^{\mathbf{j}}_{w,1}(Z_k)+\alpha[f^{\mathbf{j}}_{w,1}\star g](Z_k)}{(1-\alpha)f^{\mathbf{j}}_{w,0}(Z_k)+\alpha[f^{\mathbf{j}}_{w,0}\star g](Z_k)}. \tag{5.40}$$

We have

$$\begin{aligned}\mathbb{E}_{f_w}\big|\tilde{f}(x)-\mathfrak{A}w_{\pi(\mathbf{j})}\Lambda_{\mathbf{j}}(x)\big|^p &= 1_{w_{\pi(\mathbf{j})}=1}\mathbb{E}_{f^{\mathbf{j}}_{w,1}}\big|\tilde{f}(x)-\mathfrak{A}\Lambda_{\mathbf{j}}(x)\big|^p+1_{w_{\pi(\mathbf{j})}=0}\mathbb{E}_{f^{\mathbf{j}}_{w,0}}\big|\tilde{f}(x)\big|^p\\ &= \mathbb{E}_{f^{\mathbf{j}}_{w,0}}\Big(1_{w_{\pi(\mathbf{j})}=1}\mathfrak{Z}_{\mathbf{j}}(w)\big|\tilde{f}(x)-\mathfrak{A}\Lambda_{\mathbf{j}}(x)\big|^p+1_{w_{\pi(\mathbf{j})}=0}\big|\tilde{f}(x)\big|^p\Big).\end{aligned}$$

Noting that $w_{\pi(\mathbf{m})},\mathbf{m}\in\mathcal{M}$, are i.i.d. under $\mathbf{P}$, since π is a bijection, and taking into account that neither $1_{W_{\mathbf{j}}}, f^{\mathbf{j}}_{w,0}$ nor $\mathfrak{Z}_{\mathbf{j}}$ depends on $w_{\pi(\mathbf{j})}$ we obtain

$$\begin{aligned}&\mathbf{E}1_{W_{\mathbf{j}}}\Big(\mathbb{E}_{f_w}\big|\tilde{f}(x)-\mathfrak{A}w_{\pi(\mathbf{j})}\Lambda_{\mathbf{j}}(x)\big|^p\Big) = \mathbf{E}1_{W_{\mathbf{j}}}\mathbb{E}_{f^{\mathbf{j}}_{w,0}}\Big(\varsigma\mathfrak{Z}_{\mathbf{j}}(w)\big|\tilde{f}(x)-\mathfrak{A}\Lambda_{\mathbf{j}}(x)\big|^p+(1-\varsigma)\big|\tilde{f}(x)\big|^p\Big)\\ &\geq \mathbf{E}1_{W_{\mathbf{j}}}\mathbb{E}_{f^{\mathbf{j}}_{w,0}}\min\left[\varsigma\mathfrak{Z}_{\mathbf{j}}(w),1-\varsigma\right]\Big(\big|\tilde{f}(x)-\mathfrak{A}\Lambda_{\mathbf{j}}(x)\big|^p+\big|\tilde{f}(x)\big|^p\Big)\\ &\geq 2^{1-p}\mathfrak{A}^p\big|\Lambda_{\mathbf{j}}(x)\big|^p\mathbf{E}1_{W_{\mathbf{j}}}\mathbb{E}_{f^{\mathbf{j}}_{w,0}}\big(\min\left[\varsigma\mathfrak{Z}_{\mathbf{j}}(w),1-\varsigma\right]\big).\end{aligned} \tag{5.41}$$

Here we have used the trivial inequality $|a-b|^p\leq 2^{p-1}\big(|a|^p+|b|^p\big)$. Denote

$$\mathfrak{b}_{\mathbf{j}}(\varsigma)=\inf_{w\in W_{\mathbf{j}}}\mathbb{E}_{f^{\mathbf{j}}_{w,0}}\big(\min\left[\varsigma\mathfrak{Z}_{\mathbf{j}}(w),1-\varsigma\right]\big)$$

and remark that for any $\mathbf{j}\in\mathcal{M}$ and any $\varsigma\in\big(4/|\mathcal{M}|,1/2\big]$ in view of Tchebychev inequality

$$\mathbf{P}(W_{\mathbf{j}})\geq 1-\frac{(|\mathcal{M}|-1)\varsigma(1-\varsigma)}{(|\mathcal{M}|\varsigma+\varsigma-1)^2}\geq 1-\frac{|\mathcal{M}|\varsigma}{(|\mathcal{M}|\varsigma-1)^2}\geq 5/9. \tag{5.42}$$

Here we have used that $|\mathcal{M}|\varsigma \geq 4$. We obtain from (5.39), (5.41) and (5.42) for any $\varsigma \in \big(4/|\mathcal{M}|, 1/2\big]$

$$\inf_{\hat{f}} R\big(\hat{f}\big) \geq (5/9)2^{1-p}\mathfrak{A}^p \big\|\lambda\big\|_p^{dp}\bigg(\prod_{l=1}^{d}\sigma_l\bigg)\sum_{\mathbf{j}\in\mathcal{M}}\mathfrak{b}_{\mathbf{j}}(\varsigma). \tag{5.43}$$

2^0. Using the trivial equality $2(a\wedge b)=a+b-|a-b|$ we get for any $\mathbf{j}\in\mathcal{M}$ and $\varsigma \in \big(4/\mathcal{M}, 1/2\big]$ applying the Hölder inequality

$$2\mathfrak{b}_{\mathbf{j}}(\varsigma) \geq 1-\sup_{w\in W_{\mathbf{j}}}\sqrt{\mathbb{E}_{f^{\mathbf{j}}_{w,0}}\big\{\varsigma\mathfrak{Z}_{\mathbf{j}}(w)-(1-\varsigma)\big\}^2} \geq 1-\sup_{w\in W_{\mathbf{j}}}\sqrt{1-2\varsigma+\varsigma^2\mathbb{E}_{f^{\mathbf{j}}_{w,0}}\mathfrak{Z}^2_{\mathbf{j}}(w)}. \tag{5.44}$$

Since Z_k, $k=1,\dots,n$ are i.i.d. random vectors, we have in view of (5.40) for any $w\in W_{\mathbf{j}}$

$$\mathbb{E}_{f^{\mathbf{j}}_{w,0}}\mathfrak{Z}^2_{\mathbf{j}}(w) = \left\{\int_{\mathbb{R}^d}\frac{\big\{(1-\alpha)f^{\mathbf{j}}_{w,1}(x)+\alpha[f^{\mathbf{j}}_{w,1}\star g](x)\big\}^2}{(1-\alpha)f^{\mathbf{j}}_{w,0}(x)+\alpha[f^{\mathbf{j}}_{w,0}\star g](x)}\mathrm{d}x\right\}^n.$$

Since $f^{\mathbf{j}}_{w,1}(\cdot)=f^{\mathbf{j}}_{w,0}(\cdot)+\mathfrak{A}\Lambda_{\mathbf{j}}(\cdot)$, $f^{\mathbf{j}}_{w,0}$ is a probability density and $\int_{\mathbb{R}^d}\Lambda_{\mathbf{j}}=0$ in view of (5.6) we obtain

$$\mathbb{E}_{f^{\mathbf{j}}_{w,0}}\mathfrak{Z}^2_{\mathbf{j}}(w) = \left\{1+\mathfrak{A}^2\int_{\mathbb{R}^d}\frac{\big[(1-\alpha)\Lambda_{\mathbf{j}}(x)+\alpha[\Lambda_{\mathbf{j}}\star g](x)\big]^2}{(1-\alpha)f^{\mathbf{j}}_{w,0}(x)+\alpha[f^{\mathbf{j}}_{w,0}\star g](x)}\mathrm{d}x\right\}^n.$$

At last, $f^{\mathbf{j}}_{w,0}(x)\geq 2^{-1}f^{(0)}(x)$ for all $x\in\mathbb{R}^d$ in view of (5.11) if the first construction is used.

On the other hand, $f^{\mathbf{j}}_{w,0}=f_{w^{\mathbf{j}}}$, where the sequence $w^{\mathbf{j}}$ is obtained from w by replacing the coordinate $w_{\pi(\mathbf{j})}$ by zero. Let us remark that the definition of W implies $w^{\mathbf{j}}\in W$ for any $\mathbf{j}\in\mathcal{M}$ and, therefore $f^{\mathbf{j}}_{w,0}\in\big\{f_w, w\in W\big\}$ for any $w\in W$. Hence $[f^{\mathbf{j}}_{w,0}\star g](x)\geq 2^{-1}[f^{(0)}\star g](x)$ for all $x\in\mathbb{R}^d$ in view of (5.31) if the third construction is used.

Recalling that the third construction is used only if $\alpha=1$ we come to the following bound being true for both constructions.

$$\mathbb{E}_{f^{\mathbf{j}}_{w,0}}\mathfrak{Z}^2_{\mathbf{j}}(w) \leq \left\{1+2\mathfrak{A}^2\int_{\mathbb{R}^d}\frac{\big[(1-\alpha)\Lambda_{\mathbf{j}}(x)+\alpha[\Lambda_{\mathbf{j}}\star g](x)\big]^2}{(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)}\mathrm{d}x\right\}^n. \tag{5.45}$$

We remark that the right hand side of the obtained inequality is independent of w.

3^0. We have

$$\begin{aligned}\int_{\mathbb{R}^d}\frac{\big[(1-\alpha)\Lambda_{\mathbf{j}}(x)+\alpha[\Lambda_{\mathbf{j}}\star g](x)\big]^2}{(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)}\mathrm{d}x &= \int_{\Gamma_\emptyset}\frac{\big[(1-\alpha)\Lambda_{\mathbf{j}}(x)+\alpha[\Lambda_{\mathbf{j}}\star g](x)\big]^2}{(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)}\mathrm{d}x \\ &+ \alpha^2\sum_{J\in\mathfrak{J}^*}\int_{\Gamma_J}\frac{[\Lambda_{\mathbf{j}}\star g]^2(x)}{(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)}\mathrm{d}x.\end{aligned} \tag{5.46}$$

Here we have used that $\Lambda_{\mathbf{j}}$ is compactly supported on $[-N/4, N/4]^d$.

Let us bound the first integral in the right hand side. In view of the third assertion of Lemma 1 and Lemma 2 there exists a universal constant T such that

$$(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)\geq TN^{-d}, \quad \forall x\in\Gamma_\emptyset.$$

It yields

$$\int_{\Gamma_\emptyset}\frac{\big[(1-\alpha)\Lambda_{\mathbf{j}}(x)+\alpha[\Lambda_{\mathbf{j}}\star g](x)\big]^2}{(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)}\mathrm{d}x\le T^{-1}N^d\big[(1-\alpha)\|\Lambda_{\mathbf{j}}\|_2+\alpha\|\Lambda_{\mathbf{j}}\star g\|_2\big]^2. \tag{5.47}$$

Recall that $\mathbf{\Lambda}(x)=\prod_{l=1}^d\lambda(x_l/\sigma_l)$ and note that $\|\Lambda_{\mathbf{j}}\|_2=\|\mathbf{\Lambda}\|_2$ and $\|\Lambda_{\mathbf{j}}\star g\|_2=\|\mathbf{\Lambda}\star g\|_2$ whatever the value of $\mathbf{j}$. It yields in particular in view of (5.47)

$$\int_{\Gamma_\emptyset}\frac{\big[(1-\alpha)\Lambda_{\mathbf{j}}(x)+\alpha[\Lambda_{\mathbf{j}}\star g](x)\big]^2}{(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)}\mathrm{d}x\le T^{-1}N^d\big[(1-\alpha)\|\mathbf{\Lambda}\|_2+\alpha\|\mathbf{\Lambda}\star g\|_2\big]^2, \tag{5.48}$$

and, we remark that the latter bound is independent of $\mathbf{j}$.

In view of the Young inequality $\|\Lambda\star g\|_2\le\|\Lambda\|_2$. Thus we have

$$2\mathfrak{A}^2\int_{\Gamma_\emptyset}\frac{\big[(1-\alpha)\Lambda_{\mathbf{j}}(x)+\alpha[\Lambda_{\mathbf{j}}\star g](x)\big]^2}{(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)}\mathrm{d}x\le 2T^{-1}\|\lambda\|_2^{2d}\mathfrak{A}^2N^d\Big(\prod_{l=1}^d\sigma_l\Big). \tag{5.49}$$

If $\alpha=1$, then we need a sharper upper bound than the one in (5.49). Remind that $\check{\lambda}$ and $\check{\mathbf{\Lambda}}$ denote the Fourier transform of λ and $\mathbf{\Lambda}$ respectively. Then in view of the Plancherel theorem one has in view of Assumption 2

$$\|\Lambda\star g\|_2=(2\pi)^{-d}\int_{\mathbb{R}^d}|\check{g}(t)|^2\prod_{l=1}^d\sigma_l^2\big|\check{\lambda}(t_l\sigma_l)\big|^2\mathrm{d}t\le\Upsilon^2\prod_{l=1}^d(2\pi)^{-1}\sigma_l^2\int_{\mathbb{R}}\big|\check{\lambda}(v\sigma_l)\big|^2(1+v^2)^{-\mu_j}\mathrm{d}v.$$

Combining the obtained inequality with (5.49), we remark that the two bounds can be written in an unified way. Indeed, for any $\alpha\in[0,1]$

$$\begin{aligned}2\mathfrak{A}^2\int_{\Gamma_\emptyset}\frac{\big[(1-\alpha)\Lambda_{\mathbf{j}}(x)+\alpha[\Lambda_{\mathbf{j}}\star g](x)\big]^2}{(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)}\mathrm{d}x &\le C_{3,1}\mathfrak{A}^2N^d\prod_{l=1}^d\sigma_l^2\int_{\mathbb{R}}\big|\check{\lambda}(v\sigma_l)\big|^2|v|^{-2\boldsymbol{\mu}_l(\alpha)}\mathrm{d}v\\ &= C_{3,1}\mathfrak{A}^2N^dI^d(\alpha)\prod_{j=1}^d\sigma_l^{1+2\boldsymbol{\mu}_l(\alpha)}\end{aligned} \tag{5.50}$$

where $C_{3,1}$ is a universal constant and $I(\alpha)=\sup_{l\in\{1,\dots,d\}}\int_{\mathbb{R}}\check{\lambda}^2(u)|u|^{-2\boldsymbol{\mu}_l(\alpha)}\mathrm{d}u$. Note that $I(\alpha)$ is well defined since the conditions (5.2) and (5.6) imply

$$\big|\check{\lambda}(v)\big|\le C_{3,2}|v|^{\mathfrak{n}},$$

where $C_{3,2}$ is a constant completely determined by the function λ and by the number $\mathfrak{n}$. It yields

$$I(\alpha)\ \le\ C_{3,2}^2\int_{-1}^1|v|^{2\mathfrak{n}-2\sup_{l\in\{1,\dots,d\}}\boldsymbol{\mu}_l(\alpha)}\mathrm{d}v+\|\check{\lambda}\|_2^2=:C_{3,3}^2<\infty$$

since $2\mathfrak{n}>2\boldsymbol{\mu}_l(\alpha)$ for any $l=1,\dots,d$. We deduce from (5.50)

$$2\mathfrak{A}^2\int_{\Gamma_\emptyset}\frac{\big[(1-\alpha)\Lambda_{\mathbf{j}}(x)+\alpha[\Lambda_{\mathbf{j}}\star g](x)\big]^2}{(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)}\mathrm{d}x\le C_{3,4}\mathfrak{A}^2N^d\prod_{l=1}^d\sigma_l^{2\boldsymbol{\mu}_l(\alpha)+1}. \tag{5.51}$$

4^0. In view of the third assertion of Lemma 1 and Lemma 2 we also have

$$(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)\geq T^{-1}N^{|J|-d}\prod_{j\in J}x_j^{-2},\quad \forall x\in\Gamma_J.$$

It yields for any $J\in\mathfrak{J}^*$

$$\int_{\Gamma_J}\frac{[\Lambda_{\mathbf{j}}\star g]^2(x)}{(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)}\mathrm{d}x\leq TN^{d-|J|}\int_{\Gamma_J}\bigg([\Lambda_{\mathbf{j}}\star g](x)\prod_{j\in J}x_j\bigg)^2\mathrm{d}x. \tag{5.52}$$

Continuing (5.52) and using (5.28) we obtain by changing the variables for any $J\in\mathfrak{J}^*$

$$\int_{\Gamma_J}\frac{[\Lambda_{\mathbf{j}}\star g]^2(x)}{(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)}\mathrm{d}x\leq 4TN^{d}\int_{\mathbb{R}^d}\bigg([\mathbf{\Lambda}\star g](y)\prod_{j\in J}y_j\bigg)^2\mathrm{d}y. \tag{5.53}$$

Here we have also used that $N\geq 8$. We remark that the obtained bound is independent of $\mathbf{j}$.

5^0. Denote $R_J(x)=[\mathbf{\Lambda}\star g](x)\prod_{j\in J}x_j$ and let $\check{R}_J$ be the Fourier transform of R_J.

Then, we have for any $t=(t_1,\dots,t_d)\in\mathbb{R}^d$

$$\check{R}_J(t)=i^{|J|}\mathfrak{D}^J\big[\check{g}(t)\check{\mathbf{\Lambda}}(t)\big]=i^{|J|}\mathfrak{D}^J\bigg[\check{g}(t)\prod_{l=1}^d\sigma_l\check{\lambda}\big(t_l\sigma_l\big)\bigg]. \tag{5.54}$$

Here we have used the definition of $\mathbf{\Lambda}$ and the fact that the right hand side of (5.54) is well-defined in view of Assumption 1 and the conditions (5.2) imposed on λ.

Let $\mathscr{J}$ be an arbitrary subset of $J\cup\emptyset$. First, since $\sigma_l\leq 1, l=1,\dots,d$, we get

$$\bigg|\mathfrak{D}^{\mathscr{J}}\prod_{l=1}^d\sigma_l\check{\lambda}\big(t_l\sigma_l\big)\bigg|\leq\prod_{l=1}^d\sigma_l\Big(\big|\check{\lambda}\big(t_l\sigma_l\big)\big|\vee\big|\check{\lambda}'\big(t_l\sigma_l\big)\big|\Big),\quad\forall t\in\mathbb{R}^d, \tag{5.55}$$

where $\check{\lambda}'$ is the first derivative of the function $\check{\lambda}$.

Second, in view of Assumption 1

$$\big|\mathfrak{D}^{\mathscr{J}}\check{g}(t)\big|\leq\mathfrak{d}_1,\quad\forall t\in\mathbb{R}^d,\ \text{ if }\quad\alpha\in(0,1); \tag{5.56}$$

$$\big|\mathfrak{D}^{\mathscr{J}}\check{g}(t)\big|=\big|\check{g}(t)\check{g}^{-1}(t)\mathfrak{D}^{\mathscr{J}}\check{g}(t)\big|\leq\mathfrak{d}_2\big|\check{g}(t)\big|,\quad\forall t\in\mathbb{R}^d,\ \text{ if }\quad\alpha=1. \tag{5.57}$$

It remains to note that

$$\Big|\mathfrak{D}^J\big[\check{g}(t)\check{\mathbf{\Lambda}}(t)\big]\Big|\leq 2^{|J|}\sup_{\mathscr{J}\subseteq J\cup\emptyset}\big|\mathfrak{D}^{\mathscr{J}}\check{g}(t)\big|\sup_{\mathscr{J}\subseteq J\cup\emptyset}\big|\mathfrak{D}^{\mathscr{J}}\check{\mathbf{\Lambda}}(t)\big|,$$

and we obtain from (5.54), (5.55), (5.56) and (5.57) that for any $\mathbf{j}\in\mathcal{M}$, $J\in\mathfrak{J}^*$ and $t\in\mathbb{R}^d$

$$\begin{aligned}
\big|\check{R}_J(t)\big| &\leq 2^d\mathfrak{d}_1\prod_{l=1}^d\sigma_l\Big(\big|\check{\lambda}\big(t_l\sigma_l\big)\big|\vee\big|\check{\lambda}'\big(t_l\sigma_l\big)\big|\Big), && \alpha\in(0,1);\\
\big|\check{R}_J(t)\big| &\leq 2^d\mathfrak{d}_2\big|\check{g}(t)\big|\prod_{l=1}^d\sigma_l\Big(\big|\check{\lambda}\big(t_l\sigma_l\big)\big|\vee\big|\check{\lambda}'\big(t_l\sigma_l\big)\big|\Big)\\
&\leq 2^d\Upsilon\mathfrak{d}_2\prod_{l=1}^d\sigma_l\Big(\big|\check{\lambda}\big(t_l\sigma_l\big)\big|\vee\big|\check{\lambda}'\big(t_l\sigma_l\big)\big|\Big)(1+t_l^2)^{-\frac{\mu_l}{2}}, && \alpha=1.
\end{aligned}$$

To get the last inequality we have used Assumption 2. We remark that both bounds can be rewritten in an unified way. Namely for any $\alpha \in (0,1]$

$$\big|\check{R}_J(t)\big| \quad \le \quad \mathfrak{d}\prod_{l=1}^{d}\sigma_l\Big(\big|\check{\lambda}\big(t_l\sigma_l\big)\big| \vee \big|\check{\lambda}'\big(t_l\sigma_l\big)\big|\Big)(1+t_l^2)^{-\frac{\boldsymbol{\mu}_l(\alpha)}{2}}, \tag{5.58}$$

where we have put $\mathfrak{d} = 2^d \max[\mathfrak{d}_1, \Upsilon\mathfrak{d}_2]$.

Thus, in view of the Plancherel theorem, one has

$$\int_{\mathbb{R}^d}\bigg([\boldsymbol{\Lambda}\star g](y)\prod_{j\in J}y_j\bigg)^2\mathrm{d}y = (2\pi)^{-d}\big\|\check{R}_J\big\|_2^2 \le C_{3,5}\prod_{l=1}^{d}\sigma_l^2\int_{\mathbb{R}}\Big(\big|\check{\lambda}\big(v\sigma_l\big)\big| \vee \big|\check{\lambda}'\big(v\sigma_l\big)\big|\Big)^2|v|^{-2\boldsymbol{\mu}_l(\alpha)}\mathrm{d}v.$$

Note that $\big|\check{\lambda}'(v)\big| \le C_{3,2}|v|^{\mathfrak{n}-1}$, for any $v\in\mathbb{R}$, in view of (5.2) and (5.6), and that $2(\mathfrak{n}-1) > 2\boldsymbol{\mu}_j(\alpha)$ for any $j=1,\dots,d$,. Hence repeating the computations that led to (5.51), we obtain

$$\int_{\mathbb{R}^d}\bigg([\boldsymbol{\Lambda}\star g](y)\prod_{j\in J}y_j\bigg)^2\mathrm{d}y \le C_{3,6}\prod_{l=1}^{d}\sigma_l^{2\boldsymbol{\mu}_l(\alpha)+1}.$$

It yields together with (5.53) for any $J\in\mathfrak{J}^*$

$$2\mathfrak{A}^2\int_{\Gamma_J}\frac{[\Lambda_{\mathbf{j}}\star g]^2(x)}{(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)}\mathrm{d}x \le C_{3,7}\mathfrak{A}^2N^d\prod_{l=1}^{d}\sigma_l^{2\boldsymbol{\mu}_j(\alpha)+1}, \tag{5.59}$$

where as previously $C_{3,7}$ is a universal constant.

We deduce from (5.45), (5.46), (5.51) and (5.59) that there exists a universal constant C_4 such that for any $\alpha\in[0,1]$, $\mathbf{j}\in\mathcal{M}$ and $w\in W_j$

$$\mathbb{E}_{f^{\mathbf{j}}_{w,0}}\mathfrak{Z}^2_{\mathbf{j}}(w) \le \Big(1 + C_4\mathfrak{A}^2N^d\prod_{l=1}^{d}\sigma_l^{2\boldsymbol{\mu}_l(\alpha)+1}\Big)^n. \tag{5.60}$$

Suppose now that the following restriction holds

$$C_4\mathfrak{A}^2N^d\prod_{j=1}^{d}\sigma_l^{2\boldsymbol{\mu}_j(\alpha)+1} \le n^{-1}|\ln(\varsigma)|. \tag{5.61}$$

Under this condition $2\mathfrak{b}_{\mathbf{j}}(\varsigma) \ge 1-\sqrt{1-\varsigma} \ge 2^{-1}\varsigma$ and we obtain from (5.43) and (5.34)

$$\phi_n\big(\Sigma\big) = \inf_{\hat{f}}\sup_{f\in\Sigma}\mathcal{R}_p^{(n)}[\hat{f},f] \ge C_5\mathfrak{A}\big(N^d\varsigma\big)^{\frac{1}{p}}, \tag{5.62}$$

where, remind, Σ be either $\mathbb{N}_{\vec{r},d}(\vec{\beta},\vec{L},M)\cap\mathfrak{P}\big(\mathbb{R}^d\big)$ or $\mathbb{N}_{\vec{r},d}(\vec{\beta},\vec{L})\cap\mathbb{F}_g(R)$.

5.3.2. Generic lower bound via the second construction.

The considerations in this section are very similar to the previous ones. Let $\Sigma = \mathbb{N}_{\vec{r},d}(\vec{\beta},\vec{L})\cap\mathfrak{P}\big(\mathbb{R}^d\big)$.

We have under (5.35) and (5.36) in view of (5.16) for any estimator $\hat{f}$

$$R(\hat{f}) := \sup_{f\in\Sigma} \mathbb{E}_f \|\hat{f} - f\|_p^p \geq \mathbf{E}1_W \mathbb{E}_{f_w} \|\hat{f} - f_w\|_p^p.$$

Set for brevity $\mathfrak{B} = \|\lambda\|_1^d \mathfrak{A} \prod_{l=1}^d \sigma_l$ and remind that $w^{\mathbf{j}}$ is obtained from w by replacing the coordinate $w_{\pi(\mathbf{j})}$ by zero. For any $\mathbf{j} \in \mathcal{M}$ put $S_{\mathbf{j}} = \mathfrak{A}\Lambda_{\mathbf{j}} - \mathfrak{B} f^{(0)}$ and introduce

$$f_{w,0}^{\mathbf{j}}(\cdot) = f^{(0)}\big(1 - \mathfrak{B}S(w^{\mathbf{j}})\big) + \mathfrak{A} \sum_{\mathbf{m}\in\mathcal{M},\mathbf{m}\neq\mathbf{j}} w_{\pi(\mathbf{m})}\Lambda_{\mathbf{m}}(\cdot), \qquad f_{w,1}^{\mathbf{j}}(\cdot) = f_{w,0}^{\mathbf{j}}(\cdot) + S_{\mathbf{j}}(\cdot).$$

Denote $\tilde{f} = \hat{f} - \big(1 - \mathfrak{B}S(w^{\mathbf{j}})\big) f^{(0)}$ and remark that similarly to (5.39)

$$R(\hat{f}) \geq \sum_{\mathbf{j}\in\mathcal{M}} \int_{\Pi_{\mathbf{j}}} \mathbf{E}1_{W_{\mathbf{j}}} \Big(\mathbb{E}_{f_w} \big|\tilde{f}(x) - w_{\pi(\mathbf{j})} S_{\mathbf{j}}(x)\big|^p\Big) \mathrm{d}x \qquad . \tag{5.63}$$

Let us remark first that $f_w = 1_{w_{\pi(\mathbf{j})}=1} f_{w,1}^{\mathbf{j}} + 1_{w_{\pi(\mathbf{j})}=0} f_{w,0}^{\mathbf{j}}$ and

$$f_{w,0}^{\mathbf{j}} = f_{w^{\mathbf{j}}}, \qquad \int_{\mathbb{R}^d} S_{\mathbf{j}}(x)\mathrm{d}x = 0, \quad \forall \mathbf{j} \in \mathcal{M}. \tag{5.64}$$

Next, noting that $\|f^{(0)}\|_\infty \leq (2N)^{-d}$ and using the obvious inequality $|a-b|^p \geq 2^{1-p}|a|^p - |b|^p$, we get

$$\int_{\Pi_{\mathbf{j}}} \big|S_{\mathbf{j}}(x)\big|^p \mathrm{d}x \geq 2^{1-p}\mathfrak{A}^p \|\lambda\|_p^{dp} \Big(\prod_{l=1}^d \sigma_l\Big) - 2^{d(2-p)} N^{-dp} \|\lambda\|_1^{dp} \mathfrak{A}^p \Big(\prod_{l=1}^d \sigma_l\Big)^{p+1} \geq 2^{-p}\mathfrak{A}^p \|\lambda\|_p^{dp} \Big(\prod_{l=1}^d \sigma_l\Big),$$

for all $N \geq C$, where C is a constant completely determined by λ and $\mathfrak{a}$. Define

$$\mathfrak{Z}_{\mathbf{j}}(w) := \frac{\mathrm{d}\mathbb{P}_{f_{w,1}^{\mathbf{j}}}}{\mathrm{d}\mathbb{P}_{f_{w,0}^{\mathbf{j}}}}\big(Z^{(n)}\big) = \prod_{k=1}^n \frac{(1-\alpha) f_{w,1}^{\mathbf{j}}(Z_k) + \alpha[f_{w,1}^{\mathbf{j}} \star g](Z_k)}{(1-\alpha) f_{w,0}^{\mathbf{j}}(Z_k) + \alpha[f_{w,0}^{\mathbf{j}} \star g](Z_k)}. \tag{5.65}$$

Completely repeating the computations that led to (5.43) and taking into account (5.44) we assert that

$$\inf_{\hat{f}} R(\hat{f}) \geq C_6 \mathfrak{A}^p \Big(\prod_{l=1}^d \sigma_l\Big) \sum_{\mathbf{j}\in\mathcal{M}} \bigg(1 - \sup_{w\in W_{\mathbf{j}}} \sqrt{1 - 2\varsigma + \varsigma^2 \mathbb{E}_{f_{w,0}^{\mathbf{j}}} \mathfrak{Z}_{\mathbf{j}}^2(w)}\bigg). \tag{5.66}$$

As previously we have

$$\mathbb{E}_{f_{w,0}^{\mathbf{j}}} \mathfrak{Z}_{\mathbf{j}}^2(w) = \bigg\{\int_{\mathbb{R}^d} \frac{\big\{(1-\alpha) f_{w,1}^{\mathbf{j}}(x) + \alpha[f_{w,1}^{\mathbf{j}} \star g](x)\big\}^2}{(1-\alpha) f_{w,0}^{\mathbf{j}}(x) + \alpha[f_{w,0}^{\mathbf{j}} \star g](x)} \mathrm{d}x\bigg\}^n.$$

We have already mentioned that the definition of the set W implies that $w^{\mathbf{j}} \in W$ for any $w \in W$ and any $\mathbf{j} \in \mathcal{M}$. Hence, in view of the first equality in (5.64) $\int f_{w,0}^{\mathbf{j}} = 1$ and $\int [f_{w,0}^{\mathbf{j}} \star g] = 1$ since g is a density. It yields together with the second relation in (5.64)

$$\mathbb{E}_{f_{w,0}^{\mathbf{j}}} \mathfrak{Z}_{\mathbf{j}}^2(w) = \bigg\{1 + \int_{\mathbb{R}^d} \frac{\big[(1-\alpha) S_{\mathbf{j}}(x) + \alpha[S_{\mathbf{j}} \star g](x)\big]^2}{(1-\alpha) f_{w,0}^{\mathbf{j}}(x) + \alpha[f_{w,0}^{\mathbf{j}} \star g](x)} \mathrm{d}x\bigg\}^n.$$

Recall that (5.11) holds under (5.36), since λ is positive, and, therefore,

$$\mathbb{E}_{f^{\mathbf{j}}_{w,0}} \mathfrak{Z}^2_{\mathbf{j}}(w) \leq \left\{1 + 2\int_{\mathbb{R}^d} \frac{\big[(1-\alpha)S_{\mathbf{j}}(x) + \alpha[S_{\mathbf{j}} \star g](x)\big]^2}{(1-\alpha)f^{(0)}(x) + \alpha[f^{(0)} \star g](x)} \mathrm{d}x\right\}^n.$$

Since $S_{\mathbf{j}} = \mathfrak{A}\Lambda_{\mathbf{j}} - \mathfrak{B} f^{(0)}$ we obtain taking into account that λ is positive function

$$\mathbb{E}_{f^{\mathbf{j}}_{w,0}} \mathfrak{Z}^2_{\mathbf{j}}(w) \leq \left\{1 + 2\mathfrak{A}^2\int_{\mathbb{R}^d} \frac{\big[(1-\alpha)\Lambda_{\mathbf{j}}(x) + \alpha[\Lambda_{\mathbf{j}} \star g](x)\big]^2}{(1-\alpha)f^{(0)}(x) + \alpha[f^{(0)} \star g](x)} \mathrm{d}x - 4\mathfrak{A}\mathfrak{B}\Big(\prod_{l=1}^d \sigma_l\Big)\|\lambda\|_1^d + 2\mathfrak{B}^2\right\}^n.$$

It remains to note $\mathfrak{A}\Big(\prod_{l=1}^d \sigma_l\Big)\|\lambda\|_1^d = \mathfrak{B}$ and, therefore,

$$\mathbb{E}_{f^{\mathbf{j}}_{w,0}} \mathfrak{Z}^2_{\mathbf{j}}(w) \leq \left\{1 + 2\mathfrak{A}^2\int_{\mathbb{R}^d} \frac{\big[(1-\alpha)\Lambda_{\mathbf{j}}(x) + \alpha[\Lambda_{\mathbf{j}} \star g](x)\big]^2}{(1-\alpha)f^{(0)}(x) + \alpha[f^{(0)} \star g](x)} \mathrm{d}x\right\}^n.$$

Since the right hand side of the latter inequality is exactly the same as in (5.45), the computations which led to the lower bound (5.62) under the restriction (5.61) remain true in the case considered here as well. Note however that the calculations led to (5.61) exploit heavily the fact that

$$\int_{\mathbb{R}} \Big(\big|\check{\lambda}(v)\big| \vee \big|\check{\lambda}'(v)\big|\Big)^2 |v|^{-2\boldsymbol{\mu}_j(\alpha)}\mathrm{d}v \leq \int_{-1}^{1} \Big(\big|\check{\lambda}(v)\big| \vee \big|\check{\lambda}'(v)\big|\Big)^2 |v|^{-2\boldsymbol{\mu}_j(\alpha)}\mathrm{d}v + \big\||\check{\lambda}| \vee |\check{\lambda}'|\big\|_2^2 < \infty.$$

It was guaranteed by the condition (5.6) imposed on λ, which is not verified now since we supposed that λ is a positive function. Hence, necessarily $|\check{\lambda}(0)| \neq 0$ and the integral above converges at 0 if and only if $\boldsymbol{\mu}_j(\alpha) < 1/2$.

Thus, if $\vec{\boldsymbol{\mu}}(\alpha) \in [0, 1/2)^d$ and (5.61) holds, we have

$$\phi_n\big(\mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}) \cap \mathfrak{P}(\mathbb{R}^d)\big) = \inf_{\hat{f}} \sup_{f \in \mathbb{N}_{\vec{r},d}(\vec{\beta},\vec{L}) \cap \mathfrak{P}\big(\mathbb{R}^d\big)} \mathcal{R}_p^{(n)}[\hat{f}, f] \geq C_5 \mathfrak{A}\big(N^d \varsigma\big)^{\frac{1}{p}}. \tag{5.67}$$

5.3.3. Conclusions.

The goal of this paragraph is to put together all the conditions found in the previous sections, in order to present the results in an unified way, and then to precise the organization of the remainder of the proofs.

Remind that the parameter N is chosen sufficiently large, and that $\varsigma \leq 2^{-1}$. In each case, we have to check five conditions. Four of them are common to every case, namely:

$$N^d\varsigma \geq 2^{3d+2}\prod_{l=1}^d \sigma_l; \tag{5.68}$$

$$\mathfrak{A}\sigma_l^{-\beta_l}\big(\varsigma N^d\big)^{1/r_l} \leq \mathbf{C_1} L_l, \quad \forall l = 1, \ldots, d; \tag{5.69}$$

$$\mathfrak{A}^2 N^d \prod_{j=1}^d \sigma_j^{2\boldsymbol{\mu}_j(\alpha)+1} \leq \mathbf{C_4} n^{-1}|\ln(\varsigma)|. \tag{5.70}$$

These conditions are found in (5.37), (5.35) and (5.61) respectively. A fourth condition is that $\sigma_l \in (0, 1)$ for each l.

Lastly there is a fifth condition, which is specific to the situation under study. A first possibility is to check (5.36), i.e.

$$\mathfrak{A}N^{2d}\varsigma \quad \le \quad \mathbf{C_3}. \tag{5.71}$$

Then we deduce from (5.62)

$$\phi_n\Big(\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L}\big)\cap\mathbb{F}_g(R)\Big) \quad \ge \quad \mathbf{C_5}\mathfrak{A}\big(N^d\varsigma\big)^{\frac{1}{p}}. \tag{5.72}$$

Moreover, if $\vec{\boldsymbol{\mu}}(\alpha)\in[0,1/2)^d$ then we deduce from (5.67)

$$\phi_n\Big(\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L}\big)\cap\mathfrak{P}\big(\mathbb{R}^d\big)\Big) \quad \ge \quad \mathbf{C_5}\mathfrak{A}\big(N^d\varsigma\big)^{\frac{1}{p}}. \tag{5.73}$$

A second possibility is to check (5.10), i.e.

$$\mathfrak{A}N^{d} \quad \le \quad \mathbf{C_2}. \tag{5.74}$$

Then we deduce from (5.62)

$$\phi_n\Big(\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L},M\big)\cap\mathfrak{P}\big(\mathbb{R}^d\big)\Big) \quad \ge \quad \mathbf{C_5}\mathfrak{A}\big(N^d\varsigma\big)^{\frac{1}{p}}. \tag{5.75}$$

Our objective now is to specify the parameters $\mathfrak{A}$, N, ς and σ_l, $l=1,\dots,d$ in order to maximize the right hand side of (5.72), (5.73) and (5.75) so that the relationships (5.68), (5.69), (5.70) and either (5.71) or (5.74) are simultaneously fulfilled.

5.4. Proof of Theorems 1 and 2. Tail and dense zones.

We note that if $\varkappa_\alpha(p)>0$ the lower bounds of asymptotics of minimax risk announced in Theorems 1 and 2 are the same. Hence in view of the obvious inclusions $\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L},M\big)\subset\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L}\big)$ and $\mathfrak{P}\big(\mathbb{R}^d\big)\subset\mathbb{F}_g, g\in\mathfrak{P}\big(\mathbb{R}^d\big)$, it suffices to consider the minimax risk over $\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L},M\big)\cap\mathfrak{P}\big(\mathbb{R}^d\big)$ only. Also remind that

$$\frac{1}{\beta(\alpha)}=\sum_{j=1}^{d}\frac{2\boldsymbol{\mu}_j(\alpha)+1}{\beta_j},\qquad \frac{1}{\omega(\alpha)}=\sum_{j=1}^{d}\frac{2\boldsymbol{\mu}_j(\alpha)+1}{\beta_j r_j},\qquad L(\alpha)=\prod_{j=1}^{d}L_j^{\frac{2\boldsymbol{\mu}_j(\alpha)+1}{\beta_j}}.$$

Later on $c_i, i=1,\dots,4$, denote the constants independent of $\vec{L}$.

Tail zone: $\varkappa_\alpha(p)>p\omega(\alpha)$**.** Choose $\varsigma=c_2$, $N=\big(L(\alpha)n^{-1}\big)^{-\frac{1}{d-d/\omega(\alpha)+d/\beta(\alpha)}}$ and

$$\sigma_l=(c_1L_l)^{-\frac{1}{\beta_l}}\big(L(\alpha)n^{-1}\big)^{\frac{1/\beta_l-1/(\beta_l r_l)}{1-1/\omega(\alpha)+1/\beta(\alpha)}},\qquad \mathfrak{A}=c_3\big(L(\alpha)n^{-1}\big)^{\frac{1}{1-1/\omega(\alpha)+1/\beta(\alpha)}}.$$

Here $c_2\le 1/2$ and we remark that $N\to\infty, n\to\infty$, which guarantees that N is large enough for all n large enough.

Note also that our choice implies $\prod_{l=1}^{d}\sigma_l\le c_1^{-\frac{1}{\beta}}L(0)^{-1}$, which guarantees that (5.68) holds for all n large enough, since $N\to\infty, n\to\infty$. Also, choosing $c_3\le\mathbf{C_2}$ we assert that (5.74) holds.

Moreover (5.69) and (5.70) become respectively

$$\begin{aligned} c_1 c_3 c_2^{1/r_l} &\leq \mathbf{C_1}; & (5.76)\\ c_1^{-\frac{1}{\beta(\alpha)}} c_3^2 &\leq \mathbf{C_4} |\ln c_2|. & (5.77) \end{aligned}$$

Note also that $\sigma_l \leq (c_1 L_0)^{-\frac{1}{\beta_l}}, l = 1, \dots, d$ for all n large enough and, therefore, choosing $c_1 \geq L_0^{-1}$ we guarantee that $\sigma_l \leq 1, l = 1, \dots d$, which was the unique requirement to the choice of this sequence. Putting finally $c_2 = c_3$ we can assert that (5.76) and (5.77) are fulfilled if c_3 is chosen small enough.

It remains to note that (5.75) yields

$$\phi_n \Big(\mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}, M\big) \cap \mathfrak{P}\big(\mathbb{R}^d\big) \Big) \geq \mathbf{C_5} \big(L(\alpha) n^{-1} \big)^{\frac{1-1/p}{1-1/\omega(\alpha)+1/\beta(\alpha)}}$$

and the assertions of Theorems 1 and 2 concerning the tail zone are established.

Dense zone: $0 < \varkappa_\alpha(p) \leq p\omega(\alpha)$**.** Choose $\varsigma = c_4^{-1}$, $N^d = c_4$ and

$$\sigma_l = (c_1 L_l)^{-\frac{1}{\beta_l}} \big(L(\alpha) n^{-1} \big)^{\frac{\beta(\alpha)/\beta_l}{2\beta(\alpha)+1}}, \qquad \mathfrak{A} = c_3 \big(L(\alpha) n^{-1} \big)^{\frac{\beta(\alpha)}{2\beta(\alpha)+1}}.$$

Here c_4 is chosen large enough. Note also that our choice implies $\prod_{l=1}^d \sigma_l \to 0, n \to \infty$, which guarantees the verification of (5.68) for all n large enough. Moreover, (5.74) is fulfilled for all n large enough since $\mathfrak{A} \to 0, n \to \infty$ and $N^d = c_4$.

Additionally, (5.69) and (5.70) become respectively

$$\begin{aligned} c_1 c_3 &\leq \mathbf{C_1};\\ c_1^{-\frac{1}{\beta(\alpha)}} c_3^2 c_4 &\leq \mathbf{C_4} \ln(2). \end{aligned}$$

Both inequalities are fulfilled if we choose c_3 small enough. It remains to note that $\sigma_l \leq 1$, for any $l = 1, \dots d$, and sufficiently large n since $\sigma_l \to 0, n \to \infty$. We deduce from (5.75)

$$\phi_n \Big(\mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}, M\big) \cap \mathfrak{P}\big(\mathbb{R}^d\big) \Big) \geq \mathbf{C_5} \big(L(\alpha) n^{-1} \big)^{\frac{\beta(\alpha)}{2\beta(\alpha)+1}}$$

and the assertions of Theorems 1 and 2 concerning the dense zone are established. ∎

It will be suitable for us to continue our considerations by proving first the assertion of Theorem 2 corresponding to the sparse zone. Its proof exploits the same condition (5.74) and the required lower bound is deduced as previously from (5.75).

5.5. Proof of Theorem 2. Sparse zone.

Sparse zone 1: $\tau(\infty) > 0$. Choose $N^d = c_4$,

$$\varsigma = \big[L(0)\big]^{-\frac{(2+1/\beta(\alpha))\omega(\alpha)}{(1-1/\omega(0))(2+1/\beta(\alpha))\omega(\alpha)+1/\beta(0)}} \big(L(\alpha) n^{-1} \ln(n) \big)^{\frac{1}{(1-1/\omega(0))(2+1/\beta(\alpha))\beta(0)+1/\omega(\alpha)}}; \tag{5.78}$$

$$\mathfrak{A} = c_3 \big[L(0)\big]^{\frac{1}{(1-1/\omega(0))(2+1/\beta(\alpha))\omega(\alpha)+1/\beta(0)}} \big(L(\alpha) n^{-1} \ln(n) \big)^{\frac{(1-1/\omega(0))\beta(0)}{(1-1/\omega(0))(2+1/\beta(\alpha))\beta(0)+1/\omega(\alpha)}}. \tag{5.79}$$

and let c_4 be chosen large enough. Let us remark that (5.74) is fulfilled for all n large enough in view of $N^d = c_4$ and $\mathfrak{A} \to 0, n \to \infty$, because $\omega(0) > 1$. Define

$$\sigma_l = L_l^{-\frac{1}{\beta_l}} \big[L(0)\big]^{\frac{(1/\beta_l)-[(2+1/\beta(\alpha))\omega(\alpha)]/(\beta_l r_l)}{(1-1/\omega(0))(2+1/\beta(\alpha))\omega(\alpha)+1/\beta(0)}} \big(L(\alpha)n^{-1}\ln n\big)^{\frac{(1/\beta_l)(1-1/\omega(0))\beta(0)+1/(\beta_l r_l)}{(1-1/\omega(0))(2+1/\beta(\alpha))\beta(0)+1/\omega(\alpha)}}. \qquad (5.80)$$

First we note that $\sigma_l < 1$ for all n large enough since $\sigma_l \to 0, n \to \infty,\ l = 1, \dots d$, in view of $\omega(0) > 1$. Next, (5.68) becomes $c_4 \geq 2^{3d+2}$ and, therefore, it is verified for c_4 large enough.

At last, (5.69) becomes $c_3 c_4^{1/r_l} \leq \mathbf{C_1}$ and it is satisfied if one chooses c_3 small enough. Moreover, for all n large enough (5.70) will be satisfied if

$$c_3^2 c_4 \leq 2^{-1}\mathbf{C_4}\big[1 - 1/\omega(0))(2 + 1/\beta(\alpha))\beta(0) + 1/\omega(\alpha)\big].$$

and, therefore, it suffices to choose c_3 small enough.

We deduce from (5.75) that

$$\phi_n\Big(\mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}, M\big) \cap \mathfrak{P}\big(\mathbb{R}^d\big)\Big) \geq \mathbf{C_7}\Big([L(0)]^{-\frac{\varkappa_\alpha(p)}{\omega(\alpha)p\tau(p)}} L(\alpha)n^{-1}\ln n\Big)^{\frac{\tau(p)}{(1-1/\omega(0))(2+1/\beta(\alpha))+1/[\omega(\alpha)\beta(0)]}}. (5.81)$$

Sparse zone 2: $\tau(\infty) \leq 0$**.** Choose $\mathfrak{A} = \mathbf{C_2} c_3 c_4^{-1}$, $N^d = c_4$ and

$$\sigma_l = (c_1 L_l)^{-\frac{1}{\beta_l}} \big(L(\alpha)n^{-1}\ln n\big)^{\frac{\omega(\alpha)}{\beta_l r_l}}, \qquad \varsigma = \big(L(\alpha)n^{-1}\ln n\big)^{\omega(\alpha)}.$$

Here c_4 is chosen large enough and $c_3 \leq 1$. First we note that (5.74) is obviously fulfilled.

Next, taking into account that $L_l \geq L_0, l = 1, \dots, d$, we assert that (5.68) holds if

$$c_4\big(L(\alpha)n^{-1}\ln n\big)^{\omega(\alpha)(1-1/\omega(0))} \geq 2^{3d+2} c_1^{-\frac{1}{\beta(0)}} L_0^{-\frac{1}{\beta(0)}}.$$

Since we consider $\omega(0) \leq 1$ the latter inequality will be fulfilled if c_4 is large enough.

Moreover, $\sigma_l \leq (c_1 L_0)^{-\frac{1}{\beta_l}} \leq 1,\ l = 1, \dots d$, if c_1 is sufficiently large. Furthermore (5.69) becomes

$$\mathbf{C_2} c_4^{1/r_l - 1} c_1 c_3 \leq \mathbf{C_1}$$

and it is satisfied if one chooses c_3 small enough.

Note at last that for all n large enough (5.70) will be fulfilled if

$$\mathbf{C_2}^2 c_3^2 c_4^{-1} c_1^{-\frac{1}{\beta(\alpha)}} \leq 2^{-1}\mathbf{C_4}\omega(\alpha),$$

and it is satisfied by choosing c_4 large enough. We deduce from (5.75) that

$$\phi_n\Big(\mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}, M\big) \cap \mathfrak{P}\big(\mathbb{R}^d\big)\Big) \geq \mathbf{C_5}\big(L(\alpha)n^{-1}\ln n\big)^{\frac{\omega(\alpha)}{p}}.$$

This together with (5.81) and (3.2) proves the assertion of the theorem corresponding to the sparse zone. ∎

5.6. Proof of Theorem 1. Sparse zone.

Sparse zone 1: $\varkappa(p) \le 0,\ \tau(p^*) > 0$**.** Here we will use the choice of parameters given in (5.78)–(5.80) and let $N^d = c_4$ as before. We have already showed that (5.68)–(5.70) are fulfilled in this case. Hence we have to check (5.71) and to verify that $\sigma_l \le 1$ for any $l = 1, \dots, d$. The following relations will be helpful for this verification.

$$(1/\beta_l)(1 - 1/\omega(0))\beta(0) + 1/(\beta_l r_l) = (1/\beta_l)\big\{\tau(p^*)\beta(0) + \big[1/r_l - 1/p^*\big]\big\} > 0, \quad \forall l = 1, \dots, d;$$

$$(1 - 1/\omega(0))(2 + 1/\beta(\alpha))\beta(0) + 1/\omega(\alpha) = \tau(p^*)(2 + 1/\beta(\alpha))\beta(0) - \big[\varkappa_\alpha(p^*)/(\omega(\alpha)p^*)\big] > 0;$$

$$\frac{(1 - 1/\omega(0))\beta(0) + 1}{(1 - 1/\omega(0))(2 + 1/\beta(\alpha))\beta(0) + 1/\omega(\alpha)} = \frac{\beta(0)\tau(p^*) + 1 - 1/p^*}{(1 - 1/\omega(0))(2 + 1/\beta(\alpha))\beta(0) + 1/\omega(\alpha)} > 0.$$

To get the first inequality we have used that $\tau(p^*) > 0$ and $p^* \ge \max_{l=1,\dots,d} r_l$, while the second one is based on $\tau(p^*) > 0$ and $\varkappa_\alpha(p^*) \le 0$. The third inequality is the consequence of the second one, $\tau(p^*) > 0$ and $p^* \ge p \ge 1$.

In view of the first and second inequalities we have $\sigma_l \to 0, n \to \infty$ and therefore, $\sigma_l \le 1, l = 1, \dots, d$, for all n large enough. Note also that (5.71) is reduced to

$$c_3 c_4^2 \big[L(0)\big]^{\frac{1-(2+1/\beta(\alpha))\omega(\alpha)}{(1-1/\omega(0))(2+1/\beta(\alpha))\omega(\alpha)+1/\beta(0)}} \big(L(\alpha)n^{-1}\ln(n)\big)^{\frac{(1-1/\omega(0))\beta(0)+1}{(1-1/\omega(0))(2+1/\beta(\alpha))\beta(0)+1/\omega(\alpha)}} \le \mathbf{C_3}.$$

In view of the third inequality above the left hand side of the latter inequality tends to zero and, therefore, it is fulfilled for all n large enough.

Thus, we deduce from (5.72) and, if $\vec{\boldsymbol{\mu}}(\alpha) \in [0, 1/2)^d$, from (5.73)

$$\begin{aligned}
\phi_n\Big(\mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}\big) \cap \mathbb{F}_g(R)\Big) &\ \ge\ \mathbf{C_7}\Big(\big[L(0)\big]^{-\frac{\varkappa_\alpha(p)}{\omega(\alpha)p\tau(p)}} L(\alpha)n^{-1}\ln n\Big)^{\frac{\tau(p)}{(1-1/\omega(0))(2+1/\beta(\alpha))+1/[\omega(\alpha)\beta(0)]}};\\
\phi_n\Big(\mathbb{N}_{\vec{r},d}\big(\vec{\beta}, \vec{L}\big) \cap \mathfrak{P}(\mathbb{R}^d)\Big) &\ \ge\ \mathbf{C_7}\Big(\big[L(0)\big]^{-\frac{\varkappa_\alpha(p)}{\omega(\alpha)p\tau(p)}} L(\alpha)n^{-1}\ln n\Big)^{\frac{\tau(p)}{(1-1/\omega(0))(2+1/\beta(\alpha))+1/[\omega(\alpha)\beta(0)]}}.
\end{aligned}$$

This, together with (3.1) complete the proof of Theorem 1 in the sparse zone 1.

Zone of inconsistency and the sparse zone 2: $\varkappa_\alpha(p) \le 0,\ \tau(p^*) \le 0$. Set

$$\varpi_n = \begin{cases} \big(L(\alpha)n^{-1}|\ln(n)|\big)^{\frac{\omega(\alpha)}{\varkappa_\alpha(p^*)}}, & \varkappa_\alpha(p^*) < 0;\\ e^{n^2}, & \varkappa_\alpha(p^*) = 0;\end{cases}$$

and note that $\varpi_n \to \infty,\ n \to \infty$. In view of the latter remark, we will assume that n is large enough, such that $\varpi_n > 1$. We start our considerations with the following remark. The case $\varkappa_\alpha(p^*) = 0$ is possible only if $p^* = p$ since $\varkappa_\alpha(\cdot)$ is strictly decreasing. Choose $N^d = c_4$,

$$\sigma_l = (c_1 L_l)^{-\frac{1}{\beta_l}} \varpi_n^{\frac{r_l - p^*}{\beta_l r_l}}, \qquad \varsigma = \varpi_n^{-p^*}, \qquad \mathfrak{A} = c_3 \varpi_n. \tag{5.82}$$

Let us remark that $\mathfrak{A} \to \infty, n \to \infty$, and, therefore, (5.74) in not satisfied anymore. We will see that (5.71), which, remind, is weaker than (5.74), is fulfilled. Note also that $\mathfrak{A} \to \infty, n \to \infty$, which means that the family of functions constructed in Section 5.2 is not uniformly bounded.

Let us start with the verification of (5.68)–(5.71) which are reduced in view of (5.82) for all n large enough to

$$
\begin{aligned}
c_4\varpi_n^{-p^*\tau(p^*)} &\geq 2^{3d+2}(c_1L_0)^{-\frac{1}{\beta(0)}}; && (5.83)\\
c_1(c_4)^{1/r_l}c_3 &\leq \mathbf{C_1}; && (5.84)\\
c_4c_3^2c_1^{-\frac{1}{\beta(\alpha)}}L^{-1}(\alpha)\varpi_n^{\frac{\varkappa_\alpha(p^*)}{\omega(\alpha)}} &\leq \mathbf{C_4}p^*n^{-1}\ln(\varpi_n). && (5.85)\\
c_3c_4^2\varpi_n^{1-p^*} &\leq \mathbf{C_3}, && (5.86)
\end{aligned}
$$

Note, additionally, that $\sigma_l \leq (c_1L_0)^{-\frac{1}{\beta_l}}$ since $r_l \leq p^*$ for any $l=1,\ldots,d$. Hence, choosing c_1 large enough we guarantee that $\sigma_l \leq 1$.

Since $\tau(p^*) \leq 0$ and $\varpi_n \geq 1$ choosing c_4 large enough we guarantee the verification of (5.83). Choosing c_3 small enough we satisfy (5.84) and (5.86) since $p^* \geq 1$. If $\varkappa_\alpha(p^*) < 0$ for all n large enough (5.85) will be fulfilled if $c_4c_3^2c_1^{-\frac{1}{\beta(\alpha)}} \leq 2^{-1}\mathbf{C_4}p^*\big|\omega(\alpha)/\varkappa_\alpha(p^*)\big|$ and, therefore it suffices to choose c_3 small enough.

At last, if $\varkappa_\alpha(p^*)=0$ (5.85) becomes $c_4c_3^2c_1^{-\frac{1}{\beta(\alpha)}}L^{-1}(\alpha) \leq \mathbf{C_4}p^*n$ and it is verified for all n large enough.

Thus, we deduce from (5.72) and, if $\vec{\boldsymbol{\mu}}(\alpha) \in [0,1/2)^d$, from (5.73)

$$
\phi_n\Big(\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L}\big)\cap\mathbb{F}_g(R)\Big) \geq \mathbf{C_7}\varpi_n^{1-p^*/p}, \quad \phi_n\Big(\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L}\big)\cap\mathfrak{P}\big(\mathbb{R}^d\big)\Big) \geq \mathbf{C_7}\varpi_n^{1-p^*/p}
$$

and, first, we remark that there is no uniformly consistent estimator if $p^*=p$.

If $p^* > p$, which implies as mentioned above that $\varkappa_\alpha(p) < 0$, we obtain

$$
\phi_n\Big(\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L}\big)\cap\mathbb{F}_g(R)\Big),\ \phi_n\Big(\mathbb{N}_{\vec{r},d}\big(\vec{\beta},\vec{L}\big)\cap\mathfrak{P}\big(\mathbb{R}^d\big)\Big) \geq \mathbf{C_7}\big(L(\alpha)n^{-1}|\ln(n)|\big)^{\frac{\omega(\alpha)(1-p^*/p)}{\varkappa_\alpha(p^*)}}.
$$

This, together with (3.1) complete the proof of Theorem 1 in the sparse zone 2. ∎

6. Proof of Theorem 3.

As it was already mentioned, the proof of this theorem has many common elements with the proof of Theorem 2. In particular we will use the first construction of the finite set of functions developed in Section 5.2 and the same choice of parameters as in Section 5.5 (the sparse zone 1). However the approach used in the proof of the generic lower bound in Section 5.3 cannot be applied to the consideration of $\mathbb{L}_\infty$-risks. The approach which will be applied here is based upon the following general bound formulated in Lemma 3 below. The statement of this lemma is a simple consequence of Theorem 2.4 from Tsybakov (2009).

Lemma 3. *Let $\mathbb{F}$ be a given set of probability densities. Assume that for any sufficiently large integer n one can find a positive real number $\mathfrak{z}_n$ and a finite subset of functions $\mathcal{F}=\big\{f^{(0)},f^{(j)},\ j\in\mathcal{J}_n\big\}\subset\mathbb{F}$ such that*

$$
\big\|f^{(i)}-f^{(j)}\big\|_p \geq 2\mathfrak{z}_n, \qquad \forall i,j\in\mathcal{J}_n\cup\{0\}:\ i\neq j; \tag{6.1}
$$

$$
\limsup_{n\to\infty}\frac{1}{|\mathcal{J}_n|^2}\sum_{j\in\mathcal{J}_n}\mathbb{E}_{f^{(0)}}\left\{\frac{\mathrm{d}\mathbb{P}_{f^{(j)}}}{\mathrm{d}\mathbb{P}_{f^{(0)}}}(X^{(n)})\right\}^2 =: C < \infty. \tag{6.2}
$$

Then for any $q \geq 1$

$$\liminf_{n\to\infty} \inf_{\tilde f} \sup_{f\in\mathbb{F}} \mathfrak{z}_n^{-1}\left(\mathbb{E}_f\big\|\tilde f - f\big\|_p^q\right)^{1/q} \geq \left(\sqrt{C}+\sqrt{C+1}\right)^{-2/q},$$

where the infimum on the left hand side is taken over all possible estimators.

6.1. Proof of the theorem.

We will apply Lemma 3 with $\mathbb{F} = \mathbb{N}_{\vec r,d}\big(\vec\beta,\vec L\big)\cap\mathfrak{P}\big(\mathbb{R}^d\big)$ and $\mathcal{F} = \big\{f^{(0)}, f_w, w \in W\big\}$, where $\big\{f^{(0)}, f_w, w\in W\big\}$ is given in the first construction of Section 5.2.

We have already proved, c.f. (5.12), that under (5.5) and (5.10)

$$\big\{f^{(0)}, f_w, w\in W\big\} \subset \mathbb{N}_{\vec r,d}\big(\vec\beta,\vec L\big)\cap\mathfrak{P}\big(\mathbb{R}^d\big). \tag{6.3}$$

Let $\tilde{\mathbf{W}} = (\tilde{\mathbf{w}}_{\mathbf{1}},\dots,\tilde{\mathbf{w}}_{|\mathcal{M}|})$ be the canonical basis of $\mathbb{R}^{|\mathcal{M}|}$. Let $\mathbf{W} = (\mathbf{w_m} := \tilde{\mathbf{w}}_{\pi(\mathbf{m})}, \mathbf{m}\in\mathcal{M})$, which is the same as $\tilde{\mathbf{W}}$ up to some reordering. Thus $\mathbf{W}$ will play the role of $\mathcal{J}_n$ in the lemma.

Let $\mathbf{0}\in\{0,1\}^{\mathcal{M}}$ denote the sequence with zero entries. With this notation $f^{(0)} = f_{\mathbf{0}}$ and we have for any $\mathbf{w},\mathbf{w}'\in\mathbf{W}$ in view of (5.3) and (5.4)

$$\|f_{\mathbf{w}} - f_{\mathbf{w}'}\|_\infty = \|F_{\mathbf{w}} - F_{\mathbf{w}'}\|_\infty = \mathfrak{A}\sup_{\mathbf{m}\in\mathcal{M}}\big|\mathbf{w}_{(\mathbf{m})} - \mathbf{w}'_{(\mathbf{m})}\big| \sup_{x\in\Pi_{\mathbf{m}}}\big|\Lambda_{\mathbf{m}}(x)\big| = \mathfrak{A}\big\|\lambda\big\|_\infty^d =: 2\mathfrak{z}_n. \tag{6.4}$$

We conclude that (6.1) is verified with $\mathfrak{z}_n = 2^{-1}\mathfrak{A}\big\|\lambda\big\|_\infty^d$. It remains to check (6.2).

As before, for any $\mathbf{w}\in\mathbf{W}$, let

$$E(\mathbf{w}) := \mathbb{E}_{f^0}\left\{\frac{\mathrm{d}\mathbb{P}_{f_{\mathbf{w}}}}{\mathrm{d}\mathbb{P}_{f^{(0)}}}(X^{(n)})\right\}^2 = \left\{\int_{\mathbb{R}^d}\frac{\big\{(1-\alpha)f_{\mathbf{w}}(x)+\alpha[f_{\mathbf{w}}\star g](x)\big\}^2}{(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)}\mathrm{d}x\right\}^n.$$

Since $f^{(0)}$ is a probability density and $\int_{\mathbb{R}^d}\Lambda_{\mathbf{m}} = 0, \mathbf{m}\in\mathcal{M}$, in view of (5.6) we obtain

$$E(\mathbf{w}) = \left\{1+\mathfrak{A}^2\int_{\mathbb{R}^d}\frac{\big\{(1-\alpha)\Lambda_{\mathbf{j}}(x)+\alpha[\Lambda_{\mathbf{j}}\star g](x)\big\}^2}{(1-\alpha)f^{(0)}(x)+\alpha[f^{(0)}\star g](x)}\mathrm{d}x\right\}^n,$$

Here $\mathbf{j}\in\mathcal{M}$ is uniquely determined from the relation $\mathbf{w_j} = 1$ since π is a bijection.

We note that up to the factor 2 the latter expression coincides with the one found in (5.45). Hence the bound (5.60) is applicable and we obtain for any $\mathbf{w}\in\mathbf{W}$

$$E(\mathbf{w}) \leq \left(1 + C_4\mathfrak{A}^2N^d\prod_{l=1}^d \sigma_l^{2\boldsymbol{\mu}_l(\alpha)+1}\right)^n.$$

Taking into account that the right hand side of the latter inequality is independent of $\mathbf{w}$ and that $|\mathcal{J}_n| = |\mathbf{W}| = |\mathcal{M}|$ we can assert that (6.2) holds with $C=1$ if we assume

$$C_4\mathfrak{A}^2N^d\prod_{l=1}^d \sigma_l^{2\boldsymbol{\mu}_l(\alpha)+1} \leq n^{-1}\ln|\mathcal{M}|. \tag{6.5}$$

Choose as previously $M_l = N(8\sigma_l)^{-1}, l = 1, \dots, d$, which yields, remind, that

$$|\mathcal{M}| = (N/8)^d \Big(\prod_{l=1}^{d} \sigma_l \Big)^{-1}.$$

Choose also $N^d = c_4$, where, remind c_4 is large enough. Then (6.5) is reduced to

$$\mathfrak{A}^2 \prod_{j=1}^{d} \sigma_l^{2\boldsymbol{\mu}_j(\alpha)+1} \leq \mathbf{C_4} n^{-1} \Big| \ln \Big(\prod_{l=1}^{d} \sigma_l \Big) \Big|. \tag{6.6}$$

Noting that $S_{\mathbf{W}} = 1$ we reduce (5.5) to

$$\mathfrak{A} \sigma_l^{-\beta_l} \Big(\prod_{j=1}^{d} \sigma_j \Big)^{1/r_l} \leq \mathbf{C_1} L_l, \quad \forall l = 1, \dots, d, \tag{6.7}$$

Let us assume now that $\tau(\infty) > 0$, and let $\mathfrak{A}$ and $\sigma_l, l = 1, \dots d$, be chosen in accordance with (5.79) and (5.80) respectively. We have already checked that $\sigma_l \leq 1, l = 1, \dots d$. Note also that (5.10) is verified as well since $\mathfrak{A} \to 0, n \to \infty$ while $N = c_4$.

The simplest algebra shows that choosing c_3 in the definition of $\mathfrak{A}$ small enough we guarantee the verification of (6.6) and (6.7). The first assertion of the theorem follows now from the statement of Lemma 3 and (6.4).

Let now $\tau(\infty) \leq 0$ and choose $\mathfrak{A} = \mathbf{C_2} c_3 c_4^{-1}$ and

$$\sigma_l = (c_1 L_l)^{-\frac{1}{\beta_l}} \big(L(\alpha) n^{-1} \ln n \big)^{\frac{\omega(\alpha)}{\beta_l r_l}}.$$

Here c_4 is chosen large enough and $c_3 \leq 1$. Then (5.10) is verified and $\sigma_l \leq (c_1 L_0)^{-\frac{1}{\beta_l}} \leq 1,\ l = 1, \dots d$, if c_1 is sufficiently large. Moreover, for all n large enough (6.6) can be reduced to

$$(\mathbf{C_2})^2 c_3^2 c_4^{-2} c_1^{-\frac{1}{\beta(\alpha)}} \leq \frac{\mathbf{C_4} \omega(\alpha)}{2\omega(0)}$$

which is checked if one chooses c_3 small enough. Additionally, (6.7) can be reduced to

$$\mathbf{C_2} c_3 c_4^{-1} c_1^{1-\frac{1}{r_l \beta(0)}} \big(L(\alpha) n^{-1} \ln n \big)^{-\frac{\tau(\infty)\omega(\alpha)}{r_l}} L^{-\frac{1}{r_l}}(0) \leq \mathbf{C_1}.$$

If $\tau(\infty) < 0$ the latter inequality holds for all n large enough. If $\tau(\infty) = 0$ we have

$$\mathbf{C_2} c_3 c_4^{-1} c_1^{1-\frac{1}{r_l \beta(0)}} L^{-\frac{1}{r_l}}(0) \leq \mathbf{C_1}.$$

It remains to recall that $L(0) \geq L_0^{\frac{1}{\beta(0)}}$ and, therefore, the above inequality holds if c_3 is small enough. The second assertion of the theorem follows now from the statement of Lemma 3 and (6.4).

∎

7. Appendix

7.1. Proof of the third assertion of Lemma 1.

Set $M = N\mathfrak{a}$ and define

$$L(z) = (2M\pi)^{-1}\int_{\mathbb{R}} \big(1+(y-z)^2\big)^{-1} 1_{[-M,M]}(y)\mathrm{d}y.$$

We have for any $z \in [-M-\mathfrak{a}, M+\mathfrak{a}]$

$$\begin{aligned} L(z) &:= (2M\pi)^{-1}\int_{\mathbb{R}} \big(1+(y-z)^2\big)^{-1} 1_{[-M,M]}(y)\mathrm{d}y = (2M\pi)^{-1}\int_{-M-z}^{M-z} \big(1+u^2\big)^{-1}\mathrm{d}u \\ &\geq (2M\pi)^{-1}\int_{\mathfrak{a}}^{M} \big(1+u^2\big)^{-1}\mathrm{d}u \geq (2M\pi)^{-1}\int_{\mathfrak{a}}^{2\mathfrak{a}} \big(1+u^2\big)^{-1}\mathrm{d}u \geq AM^{-1}, \end{aligned}$$

where we have denoted $A = \mathfrak{a}\big[2\pi\{1+4\mathfrak{a}^2\}\big]^{-1}$. Here we have used that $N \geq 2 \;\Rightarrow\; M \geq 2\mathfrak{a}$.

If $z \notin [-M-\mathfrak{a}, M+\mathfrak{a}]$, noting that $|u| \geq \mathfrak{a}$ for any $u \in [-M-z, M-z]$, we get

$$L(z) \geq \mathfrak{a}^2(2M\pi\{1+\mathfrak{a}^2\})^{-1}\int_{-M-z}^{M-z} u^{-2}\mathrm{d}u = C\big(z^2-M^2\big)^{-1},$$

where we have put $C = \mathfrak{a}^2\big[\pi\{1+\mathfrak{a}^2\}\big]^{-1}$. Thus, we obtain for any $N \geq 2$

$$L(z) \geq AM^{-1}1_{[-M-\mathfrak{a},M+\mathfrak{a}]}(z) + C\big(z^2-M^2\big)^{-1}1_{\mathbb{R}\setminus[-M-\mathfrak{a},M+\mathfrak{a}]}(z), \quad \forall z \in \mathbb{R}.$$

and, therefore,

$$aL(v\mathfrak{a}) \geq AN^{-1}1_{[-N-1,N+1]}(v) + C\mathfrak{a}^{-1}\big(v^2-N^2\big)^{-1}1_{\mathbb{R}\setminus[-N-1,N+1]}(v), \quad \forall v \in \mathbb{R}. \tag{7.1}$$

It remains to note that $f_{0,N}(x) = \prod_{j=1}^d \mathfrak{a}L(\mathfrak{a}x_j)$ and that $C\mathfrak{a}^{-1} > A$ and the second assertion of the lemma follows. ∎

7.2. Proof of Lemma 2.

Set $\mathcal{A} = [-N-2N_g-1, N+2N_g+1]$ and note that (7.1) implies

$$\mathfrak{a}L(v\mathfrak{a}) \geq C_1N^{-1}, \quad \forall v \in \mathcal{A}, \;\; \forall N \geq 2,$$

where $C_1 = A \wedge \big[2C\mathfrak{a}^{-1}(5+12N_g+4N_g^2)^{-1}\big]$.

The latter inequality together with (7.1) yields

$$\begin{aligned} 2\mathfrak{a}L(v\mathfrak{a}) &\geq \mathfrak{a}L(v\mathfrak{a})1_{\mathcal{A}}(v) + \mathfrak{a}L(v\mathfrak{a})1_{\mathbb{R}\setminus[-N-1,N+1]}(v) \\ &\geq C_1N^{-1}1_{\mathcal{A}}(v) + v^{-2}C\mathfrak{a}^{-1}1_{\mathbb{R}\setminus[-N-1,N+1]}(v), \quad \forall v \in \mathbb{R}^d. \end{aligned}$$

Since, remind, $f_{0,N}(x) = \prod_{j=1}^d \mathfrak{a}L(\mathfrak{a}x_j)$ we get, putting $C_2 = 2^{-d}[C_1 \wedge C\mathfrak{a}^{-1}]^d$,

$$f_{0,N}(x) \geq C_2\prod_{j=1}^d \Big[N^{-1}1_{\mathcal{A}}(x_j) + x_j^{-2}1_{\mathbb{R}\setminus[-1-N,N+1]}(x_j)\Big], \quad \forall x \in \mathbb{R}^d. \tag{7.2}$$

For any $J \in \mathfrak{J}$ and $\mathfrak{z} \in \mathbb{R}^d$ set $\mathfrak{z}_J = \{\mathfrak{z}_j,\ j \in J\} \in \mathbb{R}^{|J|}$ and we will write $\mathfrak{z} = \big(\mathfrak{z}_J, \mathfrak{z}_{\bar{J}}\big)$.

Let $J \in \mathfrak{J}$ be fixed. We have for any $x \in \Gamma_J$ by changing the variables

$$\begin{aligned}\big[\bar{f}_{0,N} \star g\big](x) &= \int_{\mathbb{R}^{|J|}} \int_{\mathbb{R}^{d-|J|}} g\big(y_J, [x-y]_{\bar{J}}\big) f_{0,N}\big([x-y]_J, y_{\bar{J}}\big) \mathrm{d}y \\ &\geq \int_{[-N_g,N_g]^{|J|}} \int_{\mathcal{A}^{d-|J|}} g\big(y_J, [x-y]_{\bar{J}}\big) f_{0,N}\big([x-y]_J, y_{\bar{J}}\big) \mathrm{d}y, \qquad (7.3)\end{aligned}$$

since we integrate positive functions.

Note that $y_J \in [-N_g, N_g]^{|J|}$ and $x \in \Gamma_J$ imply that $|x_j - y_j| > N+1$ for any $j \in J$. This, together with $y_{\bar{J}} \in \mathcal{A}^{d-|J|}$ yields in view of (7.2)

$$f_{0,N}\big([x-y]_J, y_{\bar{J}}\big) \geq C_2 N^{|J|-d} \prod_{j \in J} (y_j - x_j)^{-2}.$$

Since $|x_j| \geq 1$ and $|y_j| \leq N_g$ for any $j \in J$

$$(y_j - x_j)^{-2} \geq \big\{2y_j^2 + 2x_j^2)\big\}^{-1} \geq (x_j)^{-2}\big\{2y_j^2 + 2)\big\}^{-1} \geq (2N_g^2 + 2)^{-1} x_j^{-2}.$$

Let $j_1 < \cdots < j_{d-|J|}$ be the elements of $\bar{J}$. We get, continuing (7.3) and putting $C_3 = C_2(2N_g^2+2)^{-d}$

$$\begin{aligned}\big[f_{0,N} \star g\big](x) &\geq C_3 N^{|J|-d} \prod_{j \in J} x_j^{-2} \int_{[-N_g,N_g]^{|J|}} \int_{\mathcal{A}^{d-|J|}} g\big(y_J, [x-y]_{\bar{J}}\big) \mathrm{d}y \\ &= C_3 N^{|J|-d} \prod_{j \in J} x_j^{-2} \int_{[-N_g,N_g]^{|J|}} \int_{-N-2N_g-1+x_{j_1}}^{N+2N_g+1+x_{j_1}} \cdots \int_{-N-2N_g-1+x_{j_{d-|J|}}}^{N+2N_g+1+x_{j_{d-|J|}}} g(u) \mathrm{d}u \\ &\geq C_3 N^{|J|-d} \prod_{j \in J} x_j^{-2} \int_{[-N_g,N_g]^d} g(u) \mathrm{d}u = B N^{|J|-d} \prod_{j \in J} x_j^{-2}, \quad \forall x \in \Gamma_J,\end{aligned}$$

where $B = 2^{-1} C_3$. Here to get the penultimate inequality we have used the definition of Γ_J while the last equality follows from the definition of N_g.

Since the collection $\big\{\Gamma_J, J \in \mathfrak{J}\big\}$ forms a partition of $\mathbb{R}^d$ we have for any $x \in \mathbb{R}^d$

$$\big[\bar{f}_{0,N} \star g\big](x) \geq B \sum_{J \in \mathfrak{J}} \Big(N^{|J|-d} \prod_{j \in J} x_j^{-2}\Big) 1_{\Gamma_J}(x).$$

∎

References

von Bahr, B. and Esseen, C.-G. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist.* **36**, 299–303.

Butucea, C. (2004) Deconvolution of supersmooth densities with smooth noise. *Canad. J. Statist.*, **32(2)**, 181–192.

Butucea, C. and Tsybakov, A. B. (2008a) Sharp optimality in density deconvolution with dominating bias. I. *Theory Probab. Appl.*, **52(1)**, 24–39.MR1700749

Butucea, C. and Tsybakov, A. B. (2008b) Sharp optimality in density deconvolution with dominating bias. II. *Theory Probab. Appl.*, **52(2)**, 237–249.

Carroll, R. J. and Hall, P. (1988) Optimal rates of convergence for deconvolving a density. *J. Amer. Statist. Assoc.*, **83**, 1184-1186.

Comte, F., Rozenholc, Y. and Taupin, M.-L. (2006) Penalized contrast estimator for adaptive density deconvolution. *Canad. J. Statist.*, **34(3)**, 431-452.

Comte, F. and Lacour, C. (2013) Anisotropic adaptive kernel deconvolution. *Ann. Inst. H. Poincaré Probab. Statist.*, **49(2)**, 569–609.

Devroye, L. (1989) The double kernel method in density estimation. *Ann. Inst. H. Poincar Probab. Statist.*, **25(4)**, 533-580.

Donoho, D. (1995) Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Appl. Comput. Harm. Anal.*, **2**, 101-126.

Donoho, D. L., Johnstone, I. M., Kerkyacharian, G. and Picard, D. (1996). Density estimation by wavelet thresholding. *Ann. Statist.* **24**, 508–539.

Fan, J. (1991) On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Statist.*, **19(3)**, 1257–1272.

Fan, J. (1993) Adaptively local one-dimensional subproblems with application to a deconvolution problem. *Ann. Statist.*, **21(2)** 600–610.

Fan, J. and Koo, J. (2002) Wavelet deconvolution. *IEEE Trans. Inform. Theory*, **48**, 734-747.

Goldenschluger, A. (1999). On pointwise adaptive nonparametric deconvolution. *Bernoulli*, **5**, 907-925.

Goldenshluger, A. and Lepski, O.V. (2011) Bandwidth selection in kernel density estimation: oracle inequalities and adaptive minimax optimality. *Ann. Statist.*, **39**, 1608-1632.

Goldenshluger, A. and Lepski, O.V. (2014) On adaptive minimax density estimation on $\mathbb{R}^d$. *Probab. Theory Related Fields*, **159**, 479–543.

Hall, P. and Meister, A. A ridge-parameter approach to deconvolution. *Ann. Statist.*, **35(4)**, 1535–1558.

Johnstone, I.M., Kerkyacharian, G., Picard, D. and Raimondo, M. (2004) Wavelet deconvolution in a periodic setting. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **66**, 547-573.

Johnstone, I. M. and Raimondo, M. (2004) Periodic boxcar deconvolution and diophantine approximation. *Ann. Statist.*, **32(5)**.

Juditsky, A. and Lambert–Lacroix, S. (2004). On minimax density estimation on $\mathbb{R}$. *Bernoulli* **10**, 187–220.

Kalifa, J. and Mallat, S. (2003) Thresholding estimators for linear inverse problems and deconvolutions. *Ann. Statist.*, **31**, 58-109.

Kerkyacharian, G., Lepski, O. and Picard, D. (2001). Nonlinear estimation in anisotropic multi–index denoising. *Probab. Theory Related Fields* **121**, 137–170.

Lepski, O.V. (2013) Multivariate density estimation under sup-norm loss: oracle approach, adaptation and independence structure. *Ann. Statist.*, **41(2)**, 1005–1034.

Lepski, O. (2015) Adaptive estimation over anisotropic functional classes via oracle approach. *Ann. Statist.*, **43(3)**, 1178–1242.

Lounici, K. and Nickl, R. (2011) Global uniform risk bounds for wavelet deconvolution estimators. *Ann. Statist.*, **39(2)**, 201–231.

Masry, E. (1993) Strong consistency and rates for deconvolution of multivariate densities of stationary processes. *Stochastic processes and their applications*, **47(1)**, 53–74.

Meister, A. (2009) *Deconvolution problems in nonparametric statistics,* volume 193 of Lecture Notes in Statistics. Springer-Verlag, Berlin.

Nikol'skii, S. M. (1977) *Priblizhenie Funktsii Mnogikh Peremennykh i Teoremy Vlozheniya. (in Rus-*

sian). [Approximation of functions of several variables and imbedding theorems.] Second edition, revised and supplemented. Nauka, Moscow.

Pensky, M. and Vidakovic, B. (1999) Adaptive wavelet estimator for nonparametric density deconvolution. *Ann. Statist.*, **27(6)**, 2033–2053.

Rebelles, G. (2015) Structural adaptive deconvolution under L_p-losses. `arXiv:1504.06246v1`, `http://arxiv.org`.

Stefanski, L. A. (1990) Rates of convergence of some estimators in a class of deconvolution problems. *Statist. Probab. Lett.*, **9**, 229-235.

Stefanski, L. A. and Carroll, R.J. (1990) Deconvoluting kernel density estimators. *Statistics*, **21**, 1696-184.

Tsybakov, A. (2009) Introduction to Nonparametric Estimation, *Springer Series in Statistics. Springer, New York.*

Willer, T. (2006) Deconvolution in white noise with a random blurring effect, *preprint LPMA*.